

CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (1), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (1) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (11) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac protein.

SQ Sequence 239 AA;

Alignment Scores:

Pred. No.:	3.17e-105	Length:	239
Score:	1201.00	Matches:	239
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	49.10%	Indels:	0
DB:	23	Gaps:	0

US-09-939-293-1 (1-1358) x AA078447 (1-239)

OY 20 ATGGCGCTCGAAGAGTGGTGGCGGCGAGCACTTCACTTCTGAGGTACAGACAG 79
DB 1 MetLaalaLeuLysSerTrpLeuSerArgSerValThrSerPheArgTyrArgGln 20
OY 80 TGTGTGTGTCTCTGTGTGTGGCTAACTTAAAGAGCGGTCTTCTGAAATTGATAGA 139
DB 21 CysLeuCysValProValValAlaAsnPhelLysArgCysPheSerGluLeuLeuArg 40
OY 140 CCATGGCACAACAACTGTGACGATTGGCTTGGAGTACCCGTGTGCGTTCTATTGCA 199
DB 41 ProTrpHisLysThrValThrIleGlyPheGlyValThrLeuCysAlaValProIleAla 60
OY 200 CAGAAATCAGAGCGCTCATCTCCCTAGTAGTGAAGATGAGAGAGCGAGTCTTG 259
DB 61 GlnLysSerGluProHisSerLeuSerSerGluAlaLeuMetArgArgAlaValSerLeu 80
OY 260 GTAACAGATACACCTTACTTCTCTCTCTGACACACATATGCGTATGGAAGCTATT 319
DB 81 ValThrAspSerThrSerThrPheLeuSerGlnThrTrpThrAlaLeuIleGluAlaIle 100
OY 320 ACTGAATATCTAAAGCGTGTATTATCACTTAACTTCTTTACCGACAATATACAGTTTA 379
DB 101 ThrGluTyrThrLysAlaValLysThrLeuThrSerLeuTyrArgGlnTyrThrSerLeu 120
OY 380 CTTCGGAAATGATGATTTCAGAGAGAGATGAAGTGTGGCAGGTGATATAGAGCCAGA 439
DB 121 LeuGlyLysMetAsnSerGluGluGluAspGluValTrpGlnValIleIleGlyAlaArg 140
OY 440 GCTGAGATGACTTCAAAAACACCAAGAGTACTTGAAGTGAAGCACTTGGATGCTGCA 499
DB 141 AlaGluMetThrSerLysHisGlnGluTyrLeuLysLeuGluTyrThrThrPheMetThrAla 160
OY 500 GTTGCTCTTTCAGAGATGGCAGCAGAACCTCATATCAAACTGGCGCAGATCAGGCCCTCT 559
DB 161 ValGlyLeuSerGluMetAlaAlaGluValAlaLysGlnThrGlyAlaAspGlnAlaSer 180
OY 560 ATAACCCGACAGATCATCTCAGCTGGTGAACCTGCAGGTGGGAAGAGGTGCACACGCTC 619
DB 181 IleThrAlaArgAsnHisIleGlnLeuValLysLeuGlnValGluGluValHisGlnLeu 200
OY 620 TCCCGAAGACGAGAACCACTGGCAGACACAGATAGAAAGCTCCGTCAGAAAACA 679
DB 201 SerArgLysAlaGluThrLysLeuAlaGluAlaGlnIleGluGluLeuArgGlnLysThr 220
OY 680 CAGGAGGAAGGGAGAGACCGGGCTGAGTGGAGCAGAGAGAGCGCTACTCCGTGAGAGAT 736
DB 221 GlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 239

RESULT 3

AAB54139
ID AAB54139 standard; Protein: 227 AA.

AC AAB54139;

DT 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:591.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neotrophic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.

KW Homo sapiens.

PN WO200055320-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05989.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WP1: 2000-579444/54.

DR N-PSDB; AAC98904.

PT New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
PS Claim 11; Page 1027-1028; 1379pp; English.

CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, neotrophic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.

SQ Sequence 227 AA;

Alignment Scores:

Pred. No.:	2.98e-99	Length:	227
Score:	1138.00	Matches:	226
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	46.52%	Indels:	0
DB:	21	Gaps:	0

US-09-939-293-1 (1-1358) x AAB54139 (1-227) ✓

OY 59 TCATCTTCAGGTACAGACAGAGTGTGGTGTGGTGGTAACTTTAAGAGCGG 118
|||||
Db 2 SerpHebhearglytrnglncylsleucysvalProvalValAlaAsnPhelylsystrg 21
OY 119 TGTTCACAGATTGATAGACATGGCACAAACTGACAGATGGCTGGTGGATACC 178
|||||
Db 22 CysHeserlgulLeuileargProtrPhisylstrValThrIlelglyPhesglyValIthr 41
OY 179 CTGTGTGGGTCTCTATTCACAGAAATCAGAGCCTCATTCCTTACGTAGTGAAGCATG 238
|||||
Db 42 LeucysAlaValProIleAlaGlnlySserGluProHissSerleuSerserGlnAlaLeu 61
OY 239 ATGAGAGAGCAGTGTCTTTGTTAAGATGACACCTTACCTTCTCTCAGACACA 298
|||||
Db 62 MetArgAlaValAlaSerleuValIthrasperThrSerThrPheleuSerGlnThrThr 81
OY 299 TATCGTTCGATTCAGCTATTTACTGAATATATACAGCTGTTTATACCTTACTCTT 358
|||||
Db 82 TyrlAlaLeuilegluAlaIlethrGluItyrThrlysalAlaValItyrThrleuThrSerleu 101
OY 359 TACCGACAATATACAGTTTACTGGGAAAATGAATTACAGAGAGAGATGAAGTGG 418
|||||
Db 102 TyrlArgIlnItyrThrSerleuLeuglyLysMetAsnSerGluGluAspGluValItrp 121
OY 419 CAGGTGATCAGAGCAGAGCTGATGATGACTCAAAACCCCAAGAGTACTTGACCTG 478
|||||
Db 122 GlnValIleIlegllyAlaAlaGlnlMetThrSerlyshISglnIutyrLeuLysleu 141
OY 479 GAAACCACTTGATGACTGAGTGTGCTTCACAGATGGCAGACAGAACTGATATCAA 538
|||||
Db 142 GlutThrThrtrpMetThrAlaValIglyLeuSerGluMetAlaAlaIleAlaItyrGln 161
OY 539 ACTGGCCAGATCAGGCTCTATACCCGACAGATACATTCAGCTGGTGAACCTGAG 598
|||||
Db 162 ThrclyAlaAspGlnAlaSerIleThrAlaArgAsnHISleGlnIleuValItyrLeuGln 181
OY 599 GTGGAAGAGGACGACGCTCTCCGGAAGACAGAAACCAAGCTGGAGAGACACAGATA 658
|||||
Db 182 ValGlnIuValIhISglnIleuSerArgLysAlaGluThrlyslleuAlaIleGlnIle 201
OY 659 GAAGAGCTCCGTCAGAAACAGAGAGAGAGGAGGAGGAGCTGAGTCGAGACAGAG 718
|||||
Db 202 GluIuLeuArgIlnIystrHnglnIuGlnIygluIuArgAlaGlnIuSerGluGlnIu 221
OY 719 GCTTACTGCTGAGAGAT 736
|||||
Db 222 AlatyrlleuArgIlnuAsp 227
|||||

RESULT 4
AAB92922

ID AAB92922 standard; Protein; 186 AA.
AC AAB92922;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11570.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length CDNA defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length CDNA -
XX
PS Claim 8; SEQ ID 11570; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length CDNA defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length CDNA. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length CDNA. The primers allow obtaining of the full-length
CC CDNA easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 186 AA;
SQ

Alignment Scores:
Pred. No.: 1,08e-74 Length: 186
Score: 879.00 Matches: 179
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.94% Indels: 0
DB: 22 Gaps: 0

US-09-939-293-1 (1-1358) x AAB92922 (1-186)

OY 200 CAGAAATCAGAGCCTTCCTTAGTGAAGCATGTGATGAGAGAGAGTGTCTTG 259
|||||
Db 8 GlnlysserGluProHissSerleuSerserGlnAlaLeuMetArgAlaValSerleu 27
OY 260 GTAACAGATGACCTTCTCTCTCAGACCAATATGGTGGATGAAGCATAT 319
|||||
Db 28 ValIthrasperThrSerThrPheleuSerGlnThrThrlyrAlaIleGlnAlaIle 47
OY 320 ACTGAATATACTAAGCTGTTTATACCTTAACCTTCTTACGACAAATATACAGTTTA 379
|||||
Db 48 ThrGluItyrThrlysalAlaValItyrThrleuThrSerleuIytrArgIlnIytrThrSerleu 67
OY 380 CTGGGAAAATGAATTGAGAGAGAGAAAGATGAGTGGCATGATCAGAGACCCAGA 439
|||||
Db 68 LeuglyLysMetAsnSerGluGlnIuAspGluValItrpGlnValIleIlelglyAlaArg 87
OY 440 GCTGAGATGACTCAAAACCAAGGACTGTAAGCTGGAACCACTTGATGACATGCA 499
|||||
Db 88 AlagluMetThrSerlyshISglnIuItyrLeuIystrleuGlnIuThrThrtrpMetThrAla 107
|||||

CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSPDYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase-3, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N39.
XX
SQ Sequence 39 AA:

Alignment Scores:
Pred. No.: 3,71e-09 Length: 39
Score: 186.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.60% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1 (1-1358) x AAU78436 (1-39)

QY 185 GCGGTTCCTATTGCACAGAAATGACAGCCTCATTCCTTACTAGTGAAGCATGTGATGAG 244

Db 1 AlavAlProIleAlaGlnLysSerGluProHisSerLeuSerSerGluAlaLeuMetArg 20

QY 245 AGACAGTGTCTTTGGTGAACAGATGACCTCTACCTTCTCTCTGACAGACATAT 301

Db 21 ArgAlaValSerLeuValThrAspSerThrSerThrPheLeuSerGlnThrThrTy 39

RESULT 7

AAU78439
ID AAU78439 standard; Peptide: 35 AA.

AC AAU78439;

DT 18-JUN-2002 (first entry)

DE Inhibitor of apoptosis (IAP) protein Smac, peptide Smac-N35.

KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1, BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;

KM neoplastic cell; tumour.

OS Homo sapiens.

XX
XX MO200216418-A2.
PN

PD 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26492.
XX
XX 24-AUG-2000; 2000US-227735P.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Alnemr1 ES;
XX
XX WPI; 2002-304115/34.
DR

PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds

PS Example 4; Page 47; 78pp; English.

XX The invention relates to an isolated Smac peptide or polypeptide (I)
XX and an isolated nucleic acid (II) encoding (I). Also described is a
XX method of identifying a compound that inhibits apoptosis, comprising:
XX (a) separately contacting several cell populations expressing a
XX cytosolic Smac (a Smac isoform that begins with MKSPDYF sequence,
XX replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
XX and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
XX domain) with a compound to be tested for apoptotic inhibiting activity;
XX (b) incubating the cell populations with a direct stimulus of the cell
XX death pathway; and (c) measuring the specific apoptotic activity of the
XX cell populations, where inhibition of the specific apoptotic activity is
XX indicative that the compound is an inhibitor of apoptosis. (I) and (II)
XX are useful for inducing apoptosis in a cell. The Smac polypeptide and
XX polynucleotide are useful for stimulating apoptosis in a neoplastic or
XX tumour cell which overexpresses an inhibitor of caspase, where the
XX inhibitor inhibits activation or activity of caspase-3, caspase-7 or
XX caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
XX (I) is useful for identifying an inhibitor or enhancer of a caspase-
XX mediated apoptosis which involves contacting a cell transformed or
XX transfected with a vector expressing (I) with a candidate inhibitor or
XX candidate enhancer; and detecting cell viability, where an increase in
XX cell viability indicates the presence of an inhibitor and a decrease in
XX cell viability indicates the presence of an enhancer. Optionally, the
XX method involves detecting the presence of large and small caspase
XX subunits after contacting cell transformed with the vector expressing
XX (I), with the candidate compound. A decrease in processing indicates the
XX presence of an inhibitor and an increase in the processing indicates the
XX presence of an enhancer. Preferably, the large and small subunits of
XX caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
XX identifying a compound that inhibits Smac binding to Smac-binding
XX molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
XX or a full-length IAP). (II) is useful in gene therapy techniques. The
XX present sequence represents the amino acid sequence of Smac peptide
XX Smac-N35.
XX

SQ Sequence 35 AA:

Alignment Scores:
Pred. No.: 4.36e-07 Length: 35
Score: 164.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.70% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1 (1-1358) x AAU78439 (1-35)

QY 185 GCGGTTCCTATTGCACAGAAATGACAGCCTCATTCCTTACTAGTGAAGCATGTGATGAG 244

Db 1 AlavAlProIleAlaGlnLysSerGluProHisSerLeuSerSerGluAlaLeuMetArg 20

QY 245 AGACAGTGTCTTTGGTGAACAGATGACCTCTACCTTCTCTCT 289

Db 21 ArgAlaValSerLeuValThrAspSerThrSerThrPheLeuSer 35

RESULT 8
 AAU78435
 ID AAU78435 standard; Peptide: 30 AA.
 AC AAU78435;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N30.
 XX
 KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
 KW neoplastic cell; mutant; tumour.
 KM
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO200216418-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 24-AUG-2001; 2001WO-US26492.
 XX
 PR 24-AUG-2000; 2000US-227735P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Alnemrl ES;
 XX
 DR WPI: 2002-304115/34.
 XX
 PT Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -
 XX
 PS Example 3: Fig 7; 78pp; English.
 XX
 CC The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MMSDPF sequence,
 CC and residues 56-60 of (I)) and an inhibitor of BID (bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the amino acid sequence of Smac mutant
 CC Smac-N30.

XX
 SO Sequence 30 AA;
 Alignment Scores:
 Pred. NO.: 6.27e-05 Length: 30
 Score: 141.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.76% Indels: 0
 DB: 23 Gaps: 0
 US-09-939-293-1 (1-1358) x AAU78435 (1-30)
 OY 185 GCGGTTCCTATTGGCAGAGAAATCAGAGCCTCATTCCTTAGTGAACATTGATGAGC 244
 Db 1 AAlaValProIleAlaGlnLysSerGluProHisSerLeuSerGluAlaLeuMetArg 20
 OY 245 AGAGCAGTGTCTTGTGTAACAGATGACCC 274
 Db 21 ArGaIaValSerLeuValThrAspSerThr 30
 RESULT 9
 AAM52315
 ID AAM52315 standard; Protein: 502 AA.
 XX
 AC AAM52315;
 XX
 DT 18-JAN-2002 (first entry)
 XX
 DE Human WASP protein.
 XX
 KW Actin polymerisation; Ena/VASP; vasodilator-stimulated phosphoprotein;
 KW metastatic cancer; parasitic infection; cytotoxic; Human; WASP.
 XX
 OS Homo sapiens.
 OS
 PN WO200171356-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-FR00843.
 XX
 PR 22-MAR-2000; 2000FR-0003637.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 PA (CURRI-) INST CURIE.
 PI Friedelz J, Friederich E, Golsteyn RM, Louvard D, Nolreaux V;
 PI Sykes C;
 XX
 DR WPI: 2001-639148/73.
 XX
 PT Identifying modulators of actin polymerization, potentially useful for
 PT treating tumor metastasis and parasitic infection, using proteins that
 PT contain Ena/VASP binding sites -
 XX
 PS Claim 13; Pages 95-96; 109pp; French.
 XX
 CC The present invention relates to a method for identifying modulators of
 CC actin polymerisation. The method involves using proteins that contain at
 CC least one binding motif for proteins of the Ena/VASP
 CC (vasodilator-stimulated phosphoprotein) family in the preparation of
 CC reagents for identification/screening of molecules that modulate
 CC formation of the actin cytoskeleton. The proteins used in the method
 CC (i.e. the proteins with binding motif(s) for Ena/VASP proteins) do not
 CC bind to the Arp2/3 protein complex. The modulators identified by the
 CC method are potentially useful for treating disorders of actin
 CC polymerisation, e.g. metastatic cancer or parasitic infection; and as
 CC cytotoxic agents. The present sequence one such protein with binding
 CC motif(s) for Ena/VASP proteins, which was used in the method of the
 CC present invention.
 XX
 SO Sequence 502 AA;


```

Db 98 -----ThrAlaArgGlyProArgArgValSerArgArgLeuProProGlnHis 113
QY 1032 CCAGGT-----AGGCAAAATGCTTTGGGTGTGAGGTAAATAATGGGTA 988
XX ||||| ||||| ||||| ||||| |||||
Db 114 ProGlyProArgGlyArgArgArgProGlyAlaGlyAlaGlyAlaProArgArgGly 133
QY 987 AGAGACCTGTACAGAGTGGGTGAATGTTAAACAGAGTGTGCCCAAGGCTAAGA 928
XX ||||| ||||| ||||| ||||| |||||
Db 134 ArgAlaArgGlyGlnAlaGlyLeu-----LeuGlyArgGlnGlyGlnGlyArg 150
QY 927 ACCAGGTCCAGGCGCAAGCTGACACACAGAGGAGCTACAGCTCAAAAGGCTTCG 868
XX ||||| ||||| ||||| ||||| |||||
Db 151 -----GlyAlaGlyArgGlyArgAlaAlaLeuGlnAlaArgArgGlyArgArg 166
QY 867 CCGATTGGCCAGGAGGAGACCTGCGCCCTCTCTGCGACAGACAGTCATGCCAAC 808
XX ||||| ||||| ||||| ||||| |||||
Db 167 Pro-----GlyProGluProAspGlnSerCys 175
QY 807 CTGGGCGAGGTGGCATCTGCCCTGCTTTCCACCTAGTGGGAGACAGGCGAGTGTGC 748
XX ||||| ||||| ||||| ||||| |||||
Db 176 GlyGlyArgProArgArgAlaAlaAlaPro-----GlyArgAlaPro--- 190
QY 747 TCAAGCCCTCAATCCTCAGCCAGGTAGGCT----- 717
XX ||||| ||||| ||||| ||||| |||||
Db 191 AlaAspProGln-ProProAlaProArgProAlaProAlaProAlaProAlaProAla 210
QY 716 -----CGCTGCCGACT---CAGCCCGCTCCTCCCTCCCTCTGTTTCTGACGG 667
XX ||||| ||||| ||||| ||||| |||||
Db 210 aaAspAlaProAlaProAlaProAlaProAlaProAlaProProProAlaSerLeu-----G 228
QY 666 AGCTCTTCTATCT-----GTGCTTCTGCGACGCTGTTCTTCTGCT 628
XX ||||| ||||| ||||| ||||| |||||
Db 228 yAlaLeuThrAlaGlySerGlyGlnGlyArgGlnSerGlnProArgAlaGlnThrLeuAr 248
QY 627 TTCCGGGAGAGCTGTGCACCTTCCACCTGCAGTTTCA-----CCAGCTGA 580
XX ||||| ||||| ||||| ||||| |||||
Db 248 gLeuGlyArg---GlyAlaProLeuProProArgAlaGlyArgGlyArgProGly 267
QY 579 ATGTAATTCCTGGGGTTATAGAGGCT 552
Db 267 nAlaGlnGlnGlnProGlyArgPro 276

```

RESULT 13

AAM40792

ID AAM40792 standard; protein; 317 AA.

XX AC AAM40792;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 5723.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

XX KW Leukemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

```

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX XX
PA (HYSE-) HYSER INC.
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
DR WPI: 2001-442253/47.
DR N-PSDB: AAI59948.
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS Example 2; SEQ ID NO 5723; 10078bp; English.
XX XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX XX
SQ Sequence 317 AA;

```

Alignment Scores:

```

Pred. No.: 0.0174 Length: 317
Score: 119.50 Matches: 82
Percent Similarity: 36.90% Conservative: 25
Best Local Similarity: 28.28% Mismatches: 99
Query Match: 4.95% Indels: 84
DB: 22 Gaps: 17

```

US-09-939-293-1 (1-1358) x AAM40792 (1-317)

QY 1320 TTGAGACGACCCAGTAAACACTGACACCCCTCTTAGTAGTAATAGTTTTCACATC 1261

Db 38 ValGlnAlaProGly-----AlaPro-----GlySer 46

QY 1260 CTTCGGCTCAGC-----TTCCTCAGAGGACGTGGGGCGAGATCAAGAAC 1213

Db 47 LeuGlyValSerHisAlaAlaAlaProProAlaArgProGlnGlyAlaAlaGlnSerPro 66

QY 1212 CATGCAAAATGATACATACAGAAAGATGTAACAACTGAGACGACCTCCCTCCCTGCCACA 1153

Db 67 HisArgGly-----ArgArgHisGlyGlyGlyAlaGlyLeuProProProArgSer 84

QY 1152 ACTGGCATCCACAGAGGAGAACAGTACTAATCATTTTTCAGACGTAATAAGACTG 1093

Db 85 ProArgPheProGlnGlnSerValProAlaSerThrSer----- 97

QY 1092 AAAACAGTTTAACAGATGCTGAACCTTAAGGCGCATGACAAAAGACTCTCTCTGAC 1033

Db 98 -----ThrAlaArgGlyProArgArgValSerArgArgLeuProProGlnHis 113

QY 1032 CCAGGT-----AGGCAAAATGCTTTGGGTGTGAGGTAAATAATGGGTA 988

Db 114 ProGlyProArgGlyArgArgArgProGlyAlaGlyAlaGlyAlaProArgArgGly 133

QY 987 AGAGACCTGTACAGAGTGGGTGAATGTTAAACAGAGTGTGCCCAAGGCTAAGA 928

Db 134 ArgAlaArgGlyGlnAlaGlyLeu-----LeuGlyArgGlnGlyGlnGlyArg 150

```
OY 927 ACCAGTCCAGCGCAGCCTAGACACAGGACCTCAGCTCACAAGGCGCTCG 868
Db 151 -----GlyAlaGluArgGluArgAlaAlaLeuGlnAlaArgGlyArg 166
OY 867 CCTGATGGCAGGCGAGCAGCTCGCGCTTCTCCGTGCACAGACAGTATGCCAAC 808
Db 167 Pro-----GlyProGluProAspGlnSerCys 175
OY 807 CTGGGCGAGGTGGATCTGCCCTTCCTTCCCATAGTGGGAGAGGAGGAGTGTGC 748
Db 176 GlyAlaArgProArgArgAlaAlaAlaPro-----GlyArgAlaPro 190
OY 747 TCAGGCGCTCAATCTCAGCAGAGTAGGCT-----717
Db 191 AlaAspProGln-ProProAlaProArgProAlaProAlaProAlaProAla 210
OY 716 -----CTGCTCCGACT---CAGCCGCTCCTCCCTCCCTGCTTTCGACGG 667
Db 210 aAspAlaProAlaProAlaProAlaProAlaProProProProHisLeu 228
OY 666 AGCTCTCTATCT-----GTGCTCTGCCAGCTTGCTTCTGCT 628
Db 228 yAlaLeuThrAlaGlySerGlyGluArgGlnSerGlnProArgAlaGluThrLeu 248
OY 627 TTCCGGAGAGCTGTGCACCTCTTCACCTGCAGTTCA-----CCAGCTGA 580
Db 248 gLeuGlyArg---GlyAlaProLeuProProArgAlaGluArgGlyArgProLys 267
OY 579 ATGTGATCTCTGGCGTTATAGAGGCT 552
Db 267 nAlaGluGlnGlnInProLysArgPro 276

RESULT 14
ABG03762
ID ABG03762 standard; Protein; 276 AA.
AC ABG03762;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3753.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS67949.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID No 34121; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
```

```
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
```

SQ Sequence 276 AA;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	0.0354	276
Percent Similarity:	116.00	65
Best Local Similarity:	35.10%	Conservative: 21
Query Match:	26.53%	Mismatches: 74
DB:	4.81%	Indels: 85
	22	Gaps: 9

US-09-939-293-1 (1-1358) x ABG03762 (1-276)

```
OY 1161 CCTGCCACAACCTGCGCATCCACAGAG-----GGAACAGTACT 1123
Db 79 ProAlaProSerGlyLeuProSerGluArgLysArgAspAlaAlaAlaLeuSer 98
OY 1122 AATTCATTTTGGACGACGTAATTAAGCTGAAAAACAGTTAAACAGTTGCTGA 1063
Db 99 AlaSerAlaLeuThr-----103
OY 1062 GGCATGACAAAAGAGACTCTCTCTGACCCAGTAGGCCAAATGCTTGGGTGAGG 1003
Db 104 GlyLeuThrLysArgProIleLeuSerSerThrProProLeuSerAlaLeuGly-----121
OY 1002 TAAAAAATGGGTAGAGCAGCTGTACAGAGTGGGTTGAATGTTAAACAGGTCAGT 943
Db 121 -----121
OY 942 GCCCAAGGCTTAAGAACAGGCTCAGCGCAAGCTGAGACACAGAGGCACTCAGAGT 883
Db 122 -----ArgLeuAlaGluAlaAlaValAla 129
OY 882 CACAAA---GGCGTCGGCGTGTGTTGGCCAGGCGAGGACTGCCGCTTCTCGGTGCA 826
Db 130 GluLysArgAlaIleSerProSerIleLysGluProSerValValProIle-----146
OY 825 CAGACAGTATGCCAAC-----CTGGCAGGAGGTGCATCTGCCCTGCTTCCCACT 772
Db 147 ---GluValLeuProThrValLeuLeuAspGluIleGluAlaAlaSerThrParAlaThr 165
OY 771 GAGTGGGAGACAGGCGAGCTGTCTCAGGCGCTCAATCTCAGCAGGAGTAGGCTCTCTGC 712
Db 166 MetThrGlySerArgAlaCysAlaGlyPro***SerSer***ArgSerProAlaPro 185
OY 711 TCCGAGTACCGCCGCTCCCTCCCTCTCTGTTTGTGACGAGCTCT-----TCT 658
Db 186 SerLeuThrAlaProSerThr***AlaSerCysThrTrpProArgSerSerProThrSer 205
OY 657 ATCTGTGCTTCCAGCAGTGTGTTCTCTCCGGAAGAGTGCAGCCTCTTCCAC 598
Db 206 SerProLeuArgAlaSerLeu-ArgLeuCysValAlaSerCysGlyGly---ThrProTr 224
OY 597 TGCAGTTTCACACGCTGAATGTGATTCCTGCGGCTTATAGAGGCTGATCTGCCAGTT 538
```


Search completed: February 20, 2003, 14:57:10
Job time : 81.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 20, 2003, 14:53:58 ; Search time 45 Seconds

(without alignments)
5802.246 Million cell updates/sec

Title: US-09-939-293-1

Perfect score: 2446

Sequence: 1 ggcgtccgcgcgcgcacaa.....ggacttaacacagaagaaaaa 1358

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_n2p.model -DEV=xlp
-O=/cgn2.1/USPTO-spool/US09939293/runat_20022003.111511.14892/app-query.fasta.1.1543
-DB=PIR.73 -QFMT=fastan -SUFIX=trp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=Pct -THR.MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09939293_eCGN_1_1_73_etrunat_20022003.111511.14892 -NCPD=6 -ICPD=3
-NO_XLIPX -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELPO=6 -DELEXT=7

Database :

PIR.73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	125.5	5.2	2	A55197
C 2	123	5.1	2	T49385
C 3	116.5	4.8	2	T22976
C 4	112	4.6	1	S02041
C 5	111	4.6	1	T7857
C 6	111	4.6	1	C45219
C 7	107.5	4.4	2	T10627
C 8	107.5	4.4	2	T23620
C 9	106.5	4.4	1	MMBEH6
C 10	105.5	4.3	2	G02520
C 11	105	4.4	2	C90968
C 12	105	4.3	2	H72552
C 13	104.5	4.3	2	T10927
C 14	104.5	4.3	2	E75383

15	104.5	4.3	1220	2	A56136	jagged protein pre
16	103.5	4.2	345	2	T08700	hypothetical prote
17	103.5	4.2	4684	2	A59404	plectin [imported]
C 18	103	4.3	420	2	T46910	hypothetical prote
19	103	4.2	592	2	B48315	lamin B2 - mouse
20	102.5	4.2	729	2	E81916	probable primosoma
21	102.5	4.2	3924	2	S37431	ankyrin 2, neuroma
22	102.5	4.2	5376	2	T42215	zonadhesin - mouse
23	102	4.2	508	1	KRSH12	keratin type II, m
24	102	4.2	2116	2	A26655	myosin heavy chain
25	101.5	4.1	373	2	T34743	hypothetical prote
26	101.5	4.1	1214	2	T00356	hypothetical prote
27	101	4.1	859	2	T26134	hypothetical prote
28	100.5	4.1	741	2	S39082	myosin heavy chain
29	100.5	4.1	1084	2	T08583	cellulose synthase
30	100	4.1	640	2	T08179	LRG5 protein - Chl
31	100	4.1	701	2	H98120	choline binding pr
32	100	4.1	4377	2	A55575	ankyrin 3, long sp
33	99.5	4.1	955	2	S24348	myosin heavy chain
34	99.5	4.1	1087	1	QFMSH	neurofilament trip
C 35	99.5	4.1	1213	2	A41724	limb deformity (ld
C 36	99	4.1	616	2	I38155	DNA-binding regula
C 37	99	4.1	660	2	JW0067	chitinase (EC 3.2.
C 38	99	4.1	721	2	E70766	hypothetical prote
C 39	99	4.1	1104	2	S59310	probable membrane
C 40	98.5	4.1	988	2	S37078	chloride channel p
41	98	4.0	548	1	QFPGI	neurofilament trip
42	98	4.0	1072	1	A37221	polyketide synthas
C 43	98	4.0	4427	2	PN0637	hypothetical prote
C 44	97.5	4.0	374	2	T16486	protein T04A6.13.1
45	97.5	4.0	762	2	G88436	

ALIGNMENTS

RESULT 1
A55197
Wiskott-Aldrich syndrome protein WASP - human
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: A54747; A55197; I38931
R:Derry, J.M.J.; Ochs, H.D.; Francke, U.
Cell 78, 635-644, 1994
A>Title: Isolation of a novel gene mutated in Wiskott-Aldrich syndrome.
A:Reference number: A54747; MUID:94349367; PMID:8069912
A:Accession: A54747
A:Molecule type: mRNA
A:Residues: 1-328, 'A', 330-366, 'LHNRPLQDLVDLHRCPLHRLVGVRCNH', 'NRHNRHRRRPARGMDFPLPSLL
A:Cross-references: GB:U12707
A:Note: this sequence is corrected in reference A55197
R:Derry, J.M.J.; Ochs, H.D.; Francke, U.
Cell 79, 922a, 1994
A:Reference number: A55197
A:Contents: erratum
A:Accession: A55197
A:Molecule type: mRNA
A:Residues: 1-424, 'PG', 427-502 <DER>
A:Cross-references: GB:U12707; NID:9695150
A:Note: the translated sequence in Genbank entry HSU12707 (PIDN:AA62663.1) differs f
Proc. Natl. Acad. Sci. U.S.A. 92, 4706-4710, 1995
A>Title: Identification of mutations in the Wiskott-Aldrich syndrome gene and charact
A:Reference number: I38931; MUID:95273432; PMID:7753869
A:Accession: I38931
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-502 <KMA>
A:Cross-references: EMBL:U19927; NID:9854672; PIDN:AAC50140.1; PID:9854673
C:Genetics:
A:Gene: GDB:WAS; IMD2; WASP
A:Cross-references: GDB:120736; OMIM:301000
A:Map position: Xp11.23-Xp11.22
A:Note: defects in this gene may result in Wiskott-Aldrich syndrome

C:Keywords: immunodeficiency

Alignment Scores:
 Pred. No.: 0.0328 Length: 502
 Score: 125.50 Matches: 74
 Percent Similarity: 35.00% Conservative: 24
 Best Local Similarity: 26.43% Mismatches: 102
 Query Match: 5.20% Indels: 80
 DB: 2 Gaps: 13

US-09-939-293-1 (1-1358) x A55197 (1-502)

```

QY 1291 CCTTCTTAGTAAATAGGTTT---TTCACTCTGGCCCTGACCTGCTCAGACGAC 1235
  ||| ::::::::::::::: ||| ||| |||
Db 205 ProaPleIethSerSerArgIyrrArgIyLeuProAlaProGlyProSerProAlaasp 224
QY 1234 -----AGTGGGGGACATCAGAGAACAACATGAAATACATAAAGAAATCGTAC 1184
  ||| ||| ||| ::::::::::::::: |||
Db 225 LysLysArgSerGlyLysLysLysLysLysLysAlaAspIleGlyAlaProSerGlyPhe 244
QY 1183 AACTGGACAGGTTCCCTCCCTCCCTGCCAC-----AACTGGCATCCCAACAGA 1136
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 Lys-----HsValSerHsValGlyTyrPaspProGlnasn 256
QY 1135 GGAACAAGTACTAATATCTTTTGACGACGTAATTAAGACTGAAGAAGAGTTAAACAGT 1076
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 Gly-----PheAspValAsnAsnLeuAspProAspLeuArgSer 269
QY 1075 TGCTAACTTAAGGACACAAAAGACTCTCTCTGTGACCCAGTAGCAAAATGC 1016
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 LeupSerArgAlaGlyLysSerIu-----AlaGlnLeu 281
QY 1015 TTTGGGTGAGGTAATAAATGGGTAAAGACAGCTGACAGAGTGGGGT----- 965
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 ThrAspAlaGluThrSerLysLeuIleTyrAspHelleGluAspGlnGlyLeuGlu 301
QY 964 -----GAAATGTTAAACAGGTGCAGTCCCA-----AGG 935
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 AlaValArgIlnuMetArgIlnuProLeuProProProProProProProSerArg 321
QY 934 GCTAAGAACCGAGTCCGACGCAAGCTGAGACACAGAGACACACAGCTCACAAGAG 875
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 GlyGlyAsnGlnLeuProArg---ProProIleValGlyGlyAsnLys----- 336
QY 874 CGTCTCGCCTGATTTGGCCAGGCGAGACCT-----GCCGCC 839
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 -----GlyArgSerGlyProLeuProProValProLeuGlyIleAlaPro 351
QY 838 TCTTCTGCTGCACACAGACAGTCATGCCAACCCCTGGGAGGGTGACATCTGCCCTTT 779
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 ProProProThrProArgGlyProProProProGlyArgGlyProProProProPro 371
QY 778 CCCCACCTAGTGGGAGAGAGGCGAGTGTCTGACGCCCCCTCAATCTCAGCAGTAGGC 719
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 ProProAlaIleHnGlyArgSerGlyProLeuProProProPro---ProGlyAlaGlyLysP 391
QY 718 CTCTGCTCGCAGTCCGCTCCCTCCCTCTCTCTGTTTCTGACGAGGAGCTCTTC 659
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 ropromerProProProProProProProProProProProSerSerGly----- 407
QY 658 TATCTGCTCTCTGCGACCTGGTTTCTGCTTTCCGAGAGACGTCGTCACCTCTTCAC 599
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 -----AsnGlyProAlaProProProProProPro 417
QY 598 CTGACGTTTACACAGCTGAAT---GTGATCTCTGGCGGTATAGAGGCTGATC 548
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 roAlaLeuValProAla-GlyGlyLeuAlaProGlyGlyGlyArgGlyAlaLeu 434

```

RESULT 2

T49385
 hypothetical protein BLD.390 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49385
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49385
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-805 <SCN>
 A:Cross-references: EMBL:AL355927; GSPDB:GN00116; NCSP:BLD1.390
 A:Experimental source: BAC clone BLD1, strain OR74A
 A:Gene: NCSP:BLD1.390
 A:Map position: 6
 A:Introns: 54/3; 212/3

Alignment Scores:
 Pred. No.: 0.0554 Length: 805
 Score: 123.00 Matches: 49
 Percent Similarity: 42.41% Conservative: 18
 Best Local Similarity: 31.01% Mismatches: 72
 Query Match: 5.10% Indels: 19
 DB: 2 Gaps: 6

US-09-939-293-1 (1-1358) x T49385 (1-805)

```

QY 927 ACCAGTCCAGGCGCAAGCTGAGACACAGGAGCATCAGAGCTCACAAGGCGCTCG 868
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 ThrArgLysSerAlaThrProThrProSerGlnThrGlyThrProGlnProProIleAla 422
QY 867 CCTGATTCGCCAGGCGAGGACCTCGCCCTCTCTCGGTGACAGACAGTCATGCCAAC 808
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 ThrProIleProProProAlaValProProValAlaGlyAlaProThrProValProAla 442
QY 807 CTGGGACAGGAGGATCTGCCCCCTGCTTCCCA-----CTGAGGGGGA 763
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 Ala-AlaAlaGlySerIleAlaAlaLeuProValIleAlaAlaValAlaLeuArgSerI 462
QY 762 GACA-----GGGCGATGTGCT-----CAGGCCCTCAATCTCAGCAGG 724
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 nThrProThrValGlyProAlaAlaProProProLeuLunlaProLysProAlaAlaSe 482
QY 723 TAGGCTCTCTGCTCCGACTCAGCCCGCTCTCCCTCTCTCTGTTTCTGACGAGC 664
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 rAlaProProThrValProAlaAlaAlaProPro---AlaProValSerAsnProSerAl 501
QY 663 TCTTCTATCTGCTTCTGCGACGCTGTTTCTGCTTCTCCGAGAGCTGGTGCACCTCT 604
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 AleuLeuAlaLeuAlaValGlnSerGlyLeuLeuSerGlyThrThrProAlaAlaProAl 521
QY 603 TCACCTGACGTTTCCACGAGCTGAATGTGAT-----TCTTGGCGGTTATAGAGGCTTGA 550
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 aProValAlaThrAlaProAlaAlaAlaProAlaAlaAlaIleThrThrTy-----Th 538
QY 549 TCTGCGCCAGTTTGTATATGACAGCTTCTGCTGCCATCTCTGAAGAAGACACAGC 498
  ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 rLeuLysGlnPheArgGlyProGlnLeuIleAsnArgLeuHsGlnLysPleu 555

```

RESULT 3

T22976
 hypothetical protein F59A2.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 A:Accession: T22976
 A:Reference number: Z19645
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1133 <NTL>
 A:Cross-references: EMBL:Z34801; PIDD:CAA84332.1; GSPDB:GN00021; CESP:F59A2.6
 A:Experimental source: clone F59A2
 R:Burton, J.

N-methyl-D-aspartate receptor chain NMDAR2D-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence (revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: C45219
R:Shih, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.; Kadotani, H.; Yokoi, M.; Akazawa
J. Biol. Chem. 268, 2836-2843, 1993
A:Title: Molecular characterization of the family of the N-methyl-D-aspartate receptor
A:Reference number: A45219; MUID:93155102; PMID:8428958
A:Accession: C45219
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1356 <1SH>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBI:124264)
C:Superfamily: N-methyl-D-aspartate receptor 2D; glutamate receptor homology
F:51-87/Domain: glutamate receptor homology <GRH>

Alignment Scores:

Pred. No.:	0.571	Length:	1356
Score:	111.00	Matches:	56
Percent Similarity:	31.37%	Conservative:	8
Best Local Similarity:	27.45%	Mismatches:	73
Query Match:	4.60%	Indels:	67
DB:	1	Gaps:	9

US-09-939-293-1 (1-1358) x C45219 (1-1356)

```
QY 915 GCAGCCTGAGCAGGAGGAGCAGCTCAGCAGCTCAGCAGGCGCTTCGCTGATGGCCA 856
    ||| :||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1004 AAlaIleAlaArgIuIngIuProthGluProProAlaIleAlaIleProGlyPhePro 1023
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 855 GGGCAGGAGCCTGCGCTTCGCTGCGACAGCAGTCAATCCAGCCGGGAGGCGG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1024 SerProProAlaProProAlaAlaAlaAlaAlaAlaAlaIleProProLeuCysArgLeu 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 795 GCA-----TCTGCCCTGCTTCCCACTGAGTGG----- 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1044 AlaPheGluAspGluSerProProAlaProSerArgTrpProArgSerAspProGluSer 1063
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 -----GGAGCAGGAGCAGTGTCTCAGCCCTCAATCCTCAGCAGGATAGCC 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1064 GluProLeuLeuGlyGlyAla-----GlyGlyProSerAlaGly-AlaProThrAl 1081
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 717 TCTGCTCCGACAGCAGCCCTCTCC-----CTTCTCCTGTTG 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1081 aProProProArgArgAlaAlaProProProCysAlaTyrLeuAspLeuGluProSerPr 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 675 TTCTGAC-----GGAGCTCTTCTATCTGTGCTTCTGCGCAGC 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1101 oSerAspSerGluAspSerGluSerLeuGlyGlyAlaSerLeuGlyGlyLeuGluProTr 1121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 639 TTGCTTCTG-----CTTCCGGGAGAGCTGCTGACCT----- 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1121 pTrpPheAlaAspPheProTrpProTyrAlaGluArgLeuGlyProProGluArgTyr 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 -----CTTCACCTGC 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1141 rTTPSerValAspLysLeuGlyGlyTrpArgAlaGlySerTrpAspTyrLeuProProAr 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 594 AGTTTACACAGT-----GAATGATTCCTGGCGTTATAGAGCGTGAATCT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1161 gGlyGlyProAlaTrpHisCysArgHisCysAlaSerLeuGluLeuLeuProProAr 1181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 GCGCCAGTTGATATGAGCTTCTGCGCATCTCTGAAGACCACTGACATCAACAA 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1181 gHis-----LeuSerCysSerHisAspLysLeuAspLysGlyTrp 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 GTGGTTTCCA 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1194 pTrpAlaPro 1197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 7
T10627
```

hypothetical protein T13K14.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence (revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T10627
R:Levan, M.; Pohl, T.; Weisenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Ler
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16991
A:Accession: T10627
A:Molecule type: DNA
A:Residues: 1-1396 <BEV>
A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.10
A:Experimental source: cultivar Columbia; BAC clone T13K14
C:Genetics:
A:Gene: ATSP:T13K14.10
A:Map position: 4
A:Insertions: 145/1; 171/2; 196/1; 242/3; 269/3; 298/3; 347/3; 374/3; 414/3; 442/3; 505,
131/3; 1139/3; 1163/3; 1194/3; 1226/3; 1259/3; 1290/3; 1344/3

Alignment Scores:

Pred. No.:	1.11	Length:	1396
Score:	107.50	Matches:	48
Percent Similarity:	36.15%	Conservative:	29
Best Local Similarity:	22.54%	Mismatches:	71
Query Match:	4.39%	Indels:	65
DB:	2	Gaps:	6

US-09-939-293-1 (1-1358) x T10627 (1-1396)

```
QY 203 AATTCAGAGCCTCATTCCTT----- 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1128 LysAspAsnProHisArgLeuValLysLeuAspAlaProGluGluAspLysAlaAla 1147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 -----AGTACTGAACATTCATGAGAGACAGTGTCTTGTGAACGATAGC 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1148 SerAlaProThrCysSerLysSerValSerGluArgLeuGluGluValArgAspThr 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 ACCTCTACCTTCCTCTCTCAGACACATATGCGTGTGATGAGCTATTACGATATAC 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1168 LysIleLysPheLeuGlyAsnLeuLysGluThrGluGluGluArgSerGluTrpArg 1187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 AAGCGTTTATACCTTACTTCTTACCGACATATACAACTTACTGGGAAATG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1188 LysLeuCysThrCysLeuLysSerGluTrpProAspTyrThrProLeuAlaLysIle 1207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 -----AATTCAGAGAGAGATGAATGATGCGAGTGATCATAGGA 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1208 LeuGluGlyLeuLeuSerArgSerAspAlaGlyAspLysIle----- 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 434 GCCAGAGCTGAGATGACTTCAAAACACAGAGTACTTGAAGTGAACCACTTGGATG 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1222 -----SerHisIleGluGluIleIleGluAlaAlaAsnGluVal 1235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 ACTGCAGTGGTCTTTTCAGATGAGAGAGCAGAGAACCTCATATCAACTGCGCAGATCAG 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1236 ArgSerValAspAlaSerGluLeu----- 1243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 GCCTCTTAACCGCAGAGATCATTCACCTGCTGAACACTGCGAGAGAGTGCAC 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1244 -----AlaArgPheLeuLeuAspLysThrIuProGluAspAspGluAlaGlu 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 CAGCTCTCCGGAAGCAGAACCAAG----- 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1260 LysLeuLysLysLysMetGluValThrArgAspGluLeuAlaAspAlaLeuTyrGluLys 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ---CTGGCAGAACGACAGATGAAGAGCTCCGTGAGAAAACACAGAGAGAGGAGAG 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1280 GlyLeuAlaMetAlaArgIleGluAsnLeuLysGlyLysGluGluGluGlu 1299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 698 CGGCTGAGTCGAGCAGAGAGGCTTACCTGCGTGAAGAT 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1300 -----GluSerSerGlnLysAspLysPheGluGluAsn 1310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 8
```



```

Db 2639 -----ProValGlyProArgAspAspPheAr 2647
QY 1054 AAAA---AGGACTCTCTCTTGACCCAGTAGGCAAAATGCTTTGGGTAGAGTAAAA 998
Db 2647 gArGLeuProSerProGlnSerSerProAlaProProAlaThrAlaProArProPr 2667
QY 997 AAAATGGGTAAAGACAGCTGTACAGAGTGGGTGAATTTAAACGGGTGACAGGCCA 938
Db 2667 oAlaSerSerArGala-----SerAlaAlaSerSer 2678
QY 937 AGGGCTAAGAACCCAGCTCAGCGCAAGCCCTGAGACCAGAGGAGCACTCAGCTCACAA 878
Db 2678 rGlySerArGalaArGArGHisArGArGAlaArGSerLeuAlaArGAlaThrGlnAlaSe 2698
QY 877 AGGCTTCGCGCTGATTTGGCCAGGCGAGAGCACTGCCCTCTTTCGTTGGTGCACAGACGT 818
Db 2698 rAlaThrThrGlnGlyTrp-----ArGProProAlaLeu---ProAspPThrVa 2713
QY 817 CATGCCAACCCCTGGGCGAGGGTG---GCATCTGCCCTTCCCTCCAGCTAGTGGGGAGA 761
Db 2713 lAlaProValThrAspPheAlaArGProProAlaProProLysProProGlnProAlaPr 2733
QY 760 CAGGCGAGTGTCTCAGGCGCTC-----AATCTCAGCGAGGTAGCGCTCTCG 713
Db 2733 oHisAlaLeuValSerGly-ValProLeuProLeuGlyProGlnAlaAlaGlyGlnAlaAs 2753
QY 712 CTCGCACATCAGCCCGCTCTCCCTCTCCCTCTGTTTCTACAGGAGCTCTTCTATCTG 653
Db 2753 ePrProAlaLeuProLLeasPrProValProProValAlaThrGly-----T 2769
QY 652 TCGCTCTGCACAGCTTGTTT-----CTGCTTCCGGAGAGAGCTGTGCAC 608
Db 2769 hValLeuProGlyLylGluAsnArGArGProProLeuThrSerGly-----ProAlaP 2787
QY 607 CTCTTCACCTGCAGCTTTCACCACTGATGTGATTCCTGGGGGTATAGAGCGCTGATC 548
Db 2787 rOthPrProArGArVal---ProValGlyGlyProGlnArGArGLeuThrArGPro---- 2804
QY 547 TCGCGCAGTTTATATGCAGCTTCGTGCGCATCTCTGAAACCAACCACTGCAGTCA---- 492
Db 2805 -----AlaValAlaSerLeuSerLLeuSerArGArGLeuP 2817
QY 491 --TCCAAGTGTTCAGAGTTCAGTCACTCTGTGTTTG----- 453
Db 2817 rOSePrProTrAspProAlaAsPrProThrAlaProValLeuGlyArGAsnProAlaGluP 2837
QY 452 --AAGTCATCTCAGCTGTGGCTCTATGATCATCTGCACACACTTCATCTTCCTCT 399
Db 2837 rOthSerSer-----SerProAlaGlyProSerProProPro 2849

RESULT 10
G02520
Plectin - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
C:Accession: G02520
R:McLean, W.H.I., Smith, F.J.D.
submitted to the EMBL Data Library, March 1996
A:Reference number: H01385
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-4574 <NCCL>
A:Cross-references: EMBL:U53204; NID:q1477645; PIDN:AB05427.1; PID:q1477646
C:Genetics:
A:Gene: PLECL1
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F;68-283/Domain: alpha-actinin actin-binding domain homology <ACT>

Alignment Scores:
Pred. No.: 1.84 Length: 4574
Score: 105.50 Matches: 38
Percent Similarity: 47.62% Conservative: 32

```

```

Best Local Similarity: 25.85% Mismatches: 48
Query Match: 4.31% Indels: 29
DB: 2 Gaps: 6
US-09-939-293-1 (1-1358) x G02520 (1-4574)
QY 386 AAAATCAATTCAGAGGAGGAGAAATGATGATGTGGCAGTATCATAGACCCAGAGCTGAG 445
Db 2073 ArGLeuGlnAlaGluGlnLysAlaHisAlaPheAlaVal-----GlnGln 2087
QY 446 ATGACTTCAAAACACCAGAGTACTTGAAGCTGGAACCACTGTGATGACTGCAGT--- 502
Db 2088 LysGluGlnGluGlnGlnGlnThrLeuGlnGlnGlnSerValLeuAspGlnLeuArG 2107
QY 503 GGTCTTTACAGATGGCAGCAGAGAGCTGCATTCAAATTCAGCGCCAGACGCTTATA 562
Db 2108 GlyGluAlaGluAlaAlaArGArGAlaAlaGluGln-----AlaGluGlnAlaArGAl 2125
QY 563 ACCGCCAGGAATCAGACTTCAGCTGTGTAACAGTGGAGAGAGGTGCACAGCTCTCC 622
Db 2126 GlnAlaGluArGluAlaAlaGlnAlaArGArGlnValGlnGlnAlaGluArGLeuLys 2145
QY 623 CGGAAAGCAGAAACCAAGCTG-----GCAGAGCACAGATAGAGAGCTC 667
Db 2146 GlnSerAlaGluGlnGlnGlnAlaGlnAlaArGAlaGlnAlaAlaAlaGluLysLeu 2165
QY 668 CGTCAGAAAACACAGAGAGAGAGGAGAGCGGCTGACTCGAGCAGAGAGGCTTACCTG 727
Db 2166 ArGlySGlnAlaGluGlnGlnAlaAlaArGArGAlaGlnAlaGluGlnAlaAla---Leu 2184
QY 728 CGTGAGGATTGAGGGGCTGAGCACACTGCCCTGTCTCCCACTCATGTGGGAAAGCAGGG 787
Db 2185 ArGln-----LysGlnAla 2189
QY 788 GCAGATGCCACCCCTGCCAGG 808
Db 2190 AlaAspAlaGluMecGluLys 2196

RESULT 11
G90968
EspF-like protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: G90968
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gesawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90968
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA86138.1; PID:q13362183; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS2715

Alignment Scores:
Pred. No.: 1.54 Length: 337
Score: 105.00 Matches: 66
Percent Similarity: 31.70% Conservative: 31
Best Local Similarity: 21.57% Mismatches: 94
Query Match: 4.35% Indels: 115
DB: 2 Gaps: 13
US-09-939-293-1 (1-1358) x G90968 (1-337)
QY 1323 CTGTTGAGAGCACAGGTAACACTCTGCACCCCTTCTTGTAGTAAATAGGTTTTC 1264
Db 55 LeuPheArGAlaProAsnAlaHisSerAlaSerPhe----- 66
QY 1263 CTCCTTGGCTCAGCTGTCTTCACAGAGACAGTGGGGCGATCAGAGAACATCAGAAA 1204

```

Db	67	-----					-----	HisArgGlnSerThrIleGlnSerLeuIleGlnGln	79
OY	1203	TAC-----					ATC	AAAGAAAT	1185
Db	80	LeuProAsnValArgGlnArgLeuIleGlnHisLeuIleGlnHisGlyLeuIle						ProAla	99
OY	1188	CGTACA-----					ACG	GCAAGTCCCGCCCGCTGC	1156
Db	100	ArgSerMetAlaGlnHisIleProProAlaProAsnTrpProAlaProProProAla							119
OY	1155	ACAACCTGGCATCCCAACAGAGGGACAACTAATCATTTTTCAGCAGTAAATAAGA							1096
Db	120	GlnAsn-----							121
OY	1095	CTGAAACACAGTTAAACAGTTGCTGAACTTAAAGGCATGACAAAGACCTCTCTCT							1036
Db	122	-----					Gln	IleSerArgProLeuPro	128
OY	1035	GACCCAGGTAGCCAAATGCTTTGGGTGCTGAGTTAAAAAAAGTTAAGACACAGCTGA							976
Db	129	AspValAla-----					Gln	ArgLeuValIleHisLeuAla	139
OY	975	CAGAGTGGGGGAATGTGTTAAACAGGGTGCATGCCCAAGGGCTAAGAACACAGTCCAGC							916
Db	140	GlnHisGlyLe-----							143
OY	915	GCAAGCCTGAGACCAACAGAGGACACTCACAGCTCACAAAGGCGCTCGCCTGATTTGCCA							856
Db	144	-----GlnProAlaArgAsnMetAlaGlnHisIleProProAlaProAsnTrpPro							160
OY	855	GGGCGAGACCTGGCCGCTTTCTCGGGTGCACACACAGTCAATGCCAACCCCTGGGCGAGGGT							796
Db	161	Ala-----							177
OY	795	GCATCGCCCTGCTTCCCACTGATGGGGAGACAGGGGACAGTGTGCTCAGGCCCTCA							736
Db	177	IleAlaGlnArgLeuValIleHisIleAsnIleGlnHisGlyIleGlnProAlaArgSerMetAla							197
OY	735	TCTCACGACGAGTAGGCGCTCTGCTCCGACTCAGCCCGCTCTCCCTCTCTCTGTGT							676
Db	197	aglnHisIle-----							214
OY	675	TTTCGACGAGACCTTCTATCTGTGGCTTCCGA-----					GCT	GGTTTCTGCTTTCCGG	622
Db	214	nasnleu-----					Gln	SerArgProLeuProAsnValAlaGlnArgLeuMetIleHis	231
OY	621	GAGACGTGGTGCACCTCTTCCACCTGACAGTTTACACAGCTGA-----					TGTGA		574
Db	231	slenAlaGlnHisGlyIleGlnProAlaArgAsnMetAlaGlnHisIleProProAlaArg							251
OY	573	TTTCTGGCGGGTTATAGAGGCTGATCTGGCCACGTTTGATATGACAGCTTCTGCTGCATC							514
Db	251	oasnTrpProAlaProThrProProVal-----					Gln	AsnGlnGlnSerArgProLeuProAsn	270
OY	513	TCTGAAGACCAACTG							498
Db	270	pvalAlaGlnArgLeu							275

RESULT 12
 H72552
 hypothetical protein APEI708 - Aeropyrum pernix (strain KL)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: H72552
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: H72552
 A:Status: preliminary
 A:Molecule type: DNA

```

A/Residues: 1-791 <RAM>
A/Cross-references: DDBJ:AP000062; NID:g5105244; PIDD:BAAB0709.1; PID:dl044495; PID:g
A/Experimental source: strain K1
C/Genetics:
A/Gene: APEI708

Alignment Scores:
Pred. No.: 1.68 Length: 791
Score: 105.00 Matches: 60
Percent Similarity: 40.51% Conservative: 36
Best Local Similarity: 25.32% Mismatches: 88
Query Match: 4.29% Indels: 53
DB: 2 Gaps: 9

```

US-09-939-293-1 (1-1358) x H72552 (1-791)

[illegible]

```

RESULT 13
110927
3C3_18c protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: 110927
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
:Reference number: 21715

```

A:Accession: T10927
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-517 <PAR>
A:Cross-references: EMBL:AL031231; NID:e1315070; PID:e1315088
A:Experimental source: strain A3(2)
C:Genetics: A:Gene: 3C3.18c

Alignment Scores:

Alignment No.:	1.77	Length:	517
Score:	104.50	Matches:	44
Percent Similarity:	37.50%	Conservative:	19
Best Local Similarity:	26.19%	Mismatches:	74
Query Match:	4.27%	Indels:	33
	2	Gaps:	5

US-09-939-293-1 (1-1358) x T10927 (1-517)

OY 424 GATCATAGAGACCGACAGCTGAGATGACTTCAAAACACCAAGAGATGACTGAAGCTGGAAC 483
Db 185 GATTCATAGAGACCGACAGCTGAGATGACTTCAAAACACCAAGAGATGACTGAAGCTGGAAC 483
OY 484 CACTTGATGATGACTGAGTGTCTTTCAGAGATGACGACAGACAGAGCTGATTCAAACTGG 543
Db 205 HSPHGGTGAATGATGAGTGTCTTTCAGAGATGACGACAGACAGAGCTGATTCAAACTGG 543
OY 544 CGCAGATGACGCTTATTAACCGCCAGGAAATCACATTCAGCTGTAAGTGAAGTGA 603
Db 212 ---LysProGlyGluTyrIleSerValProSerHisGlyAlaSerGlyAlaGly 230
OY 604 AGAGTGCACACGCTCTCCCGAAGACAGAAACCAAGCTGCGACAGAGACAGATAGAAGA 663
Db 231 SerGlyThrProAlaValAlaGlnGlyGlnAspSerGlySerArgValGlyAlaGln 250
OY 664 GCTCGTCAAGAAACACACAGAGAGAGGAGGAGCGGCTG----- 704
Db 251 ---ProHisProTyrSerGlyAlaSerGlnGlyAlaSerProGlyLeuSerSerAlaSer 269
OY 705 AGTCGAGACAGAGCGCTACTGCTGAGAGATTTGAGGCTGAGACACCTGCTG-TCT 763
Db 270 AlaAspSerArgArgProValAlaMetAspIleAspGlyLeuThrThrLeuProGlnArg 289
OY 764 CCCACCTAGTGGGAAAGCAGGGGCGAGATCCACCTGCCACG----- 808
Db 290 ProGlyProThrProGlnProGlyProAsnThrProLeuProArgSerAspAlaProPro 309
OY 809 -----GTTGGCATGACTGTCTGTGCACCGAAGAAAGCGGCGAGTCTCGCCCT 856
Db 310 GlnGlnProValValGlnGlyProProValLeuGlnAsnSerGlyGlyGlyThrAsnAsnPro 329
OY 857 GGCCATCAGGAGGAGACGCTTTG 880
Db 330 Ala---SerGlyArgMetProVal 336

RESULT 14
E75383
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75383
R:White, O.; Eelsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S. Smith, H.O.; Venter, J.C.; Fraser, C.M.
S: Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75383
A:Molecule type: DNA
A:Residues: 1-581 <NH2>
A:Cross-references: GB:AE001998; GB:AE000513; NID:96459302; PIDN:AAF1109.1; PID:96459302
A:Experimental source: strain R1

[illegible]

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1220 <LIN>
 A:Cross-references: GB:L38483
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 F:379-410/Domain: EGF homology <EGF1>
 F:492-523/Domain: EGF homology <EGF>
 F:634-665/Domain: EGF homology <EGF2>

Alignment Scores:

Pred. No.:	1.94	Length:	1220
Score:	104.50	Matches:	71
Percent Similarity:	29.67%	Conservative:	37
Best Local Similarity:	19.51%	Mismatches:	131
Query Match:	4.27%	Indels:	125
	2.	Gaps:	19

US-09-939-293-1 (1-1358) x A56136 (1-1220)

```

QY 481 AACCACTTGATGACTGCAGTTGCTCT----- 507
Db 667 AsnAsnIleAsnAspCysSerGlnAsnProCysHisTyrGlyGlyThrCysArgAspLeu 686
QY 508 TTCACAGATGGCAGCAGAAAGCTGCATATCAAACTGGCCAGATCAGCCCTATACCGC 567
Db 687 ValAsnAspPheTyrCysAspCysLysAsnGlyTyrLysGlyLysThrCysHisSerArg 706
QY 568 CAGGAATCACAATTCAGCTGTGAAACTGCAGGTGGAAGAGTCACAGCTCTCCCGAA 627
Db 707 AspSer-----GlnCysAspGluAlaThrCysAsnAsnGlyGlyThrCysTyrAspGlu 724
QY 628 -----AGCAAAACCAAGCT 642
Db 725 ValAspThrPheLysCysMetCysProGlyGlyTyrGluGlyThrThrCysAsnIleAla 744
QY 643 GGCAGAACCAACAGATAGAAGCTCCGTCAGAAAAACAGAGAGAGAGAGAGAGCGGC 702
Db 745 ArgAsnSerSerCysLeuProAsnProCysHisAsnGlyGly----- 758
QY 703 TGAATCGAGCAGAGAGCGCTACCTGCGTAGAGATTGAGGCGCTGAGCACACTGCCCTGTC 762
Db 759 -----ThrCysValVal 762
QY 763 TCCCCACTCAGTGGGGGAAAGAGGGGAGATGCCACCCGTCAGAGGTGGCAGATGACTGT 822
Db 763 -----AsnGlyAspSer---PheThrCysValCysLysGluGlyTyrGluPro 778
QY 823 CTGTGCACCGAAGAGAGCGGAGCTCTGCCCTGGCCA----- 861
Db 779 IleCysThrGlnAsnThrAsnAspCysSerProHisProCysTyrAsnSerGlyThrCys 798
QY 862 ATCAGGCGAGACGCTTGTGATGTGATGTCCTCTGTCGTCACAGCTTGGCGCTGA 921
Db 799 ValAspGlyAspAsnTyrPheArgCysGluCys-----AlaProGlyPheAlaGly 815
QY 922 CCT-----GGTTCTTAGCCCTTGGGCACTGCACCTGCATTTTAACATTTCAACCC 969
Db 816 ProAspCysArgIleAsnIleAsnGlyCysIleCysProCys---AlaPheGlyAla 834
QY 970 ACTCTG-----TACAGCTGCTCTTACCCATTTTTTTACCTCACACC 1011
Db 835 ThrCysValAspGluIleAsnGlyTyrGlnCysIleCysPro----- 848
QY 1012 CAAGACATTTTGCCTACCTGGGTCAAGAGAGAGTCTTTTGTGATGCCCTTAAGTTC 1071
Db 849 -----ProGlyHisSerGlyAlaLysCysHisGluValSerGly 861
QY 1072 AGCAACTGTTTAACCTGTTTTCAGTCTTATTTAGCTGCTCAAAAATGATTTAGTACTGT 1131
Db 862 ArgSerCysIleThrMetGlyArgValIleLeuAspGlyAlaLysTyrAspAspAspCys 881
QY 1132 TCCTCTGT-----TGGGATGCCAGTTGT 1155
Db 1132 TCCTCTGT-----TGGGATGCCAGTTGT 1155

```

```

Db 882 AsnThrCysGlnCysLeuAsnGlyArgValAlaCysSerLysValTyr-----Cys 898
QY 1156 GGC-----AGGGGAGGGGAACTGTGCCA----- 1179
Db 899 GlyProAlaGProCysArgLeuHisLysGlyHisGlyGluCysProAsnGlyGlnSerCys 918
QY 1180 GTTGTACGATTTCTTTGTATGTATTTCTGATGTGTTCTCTGATCTGGCCCACTGT--- 1236
Db 919 IleProValLeuAspAspGlnCysPheValAlaProCysThrGlyAlaGlyGluCysArg 938
QY 1237 -----CCTGTAGAGACAGCTAGGCCAAGAGTGAATAAATCTATTACTAC 1281
Db 939 SerSerSerLeuGlnProValLysThrLysCysThrSerAsp-----SerTyrTyr--- 955
QY 1282 TAAGAGAGGGGTGCAGAGTGTTTACCTGCTGCTCTCAACAGAGACTTAAATCAACAGGA 1341
Db 956 ---GlnAspAsnCysAlaAsnIleThrPheThrPheAsnLysGluMetSerProGly 974
QY 1342 CTTAACACAGAA 1353
Db 975 LeuThrThrGlu 978

```

Search completed: February 20, 2003, 15:03:02
 Job time : 69 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 20, 2003, 14:54:33 : Search time 18.5 Seconds

(without alignments)
4319.603 Million cell updates/sec

Title: US-09-939-293-1

Perfect score: 2446
Sequence: 1 ggcgtccggcgctgcacaa.....ggacttaacacagaaaaaa 1358

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+.n2p.model -DEV-x1p
-O=/cgn2_1/USPTO.spool/US0939293/runat_20022003_111511_14911/app_query.fasta-1.1543
-DB-issued_patents_AA -OPMT-fastan -SUFFIX-ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US0939293_6CGN_1_1_13_etunal_20022003_111511_14911 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1201	49.1	239	3	US-09-479-309-2
2	106	4.3	591	3	US-08-965-903B-2
3	105.5	4.4	498	2	US-08-660-963-12
4	104.5	4.3	1219	4	US-08-882-046-5
5	100	4.1	564	4	US-09-308-022-6
6	98	4.1	1214	2	US-08-231-193A-54
7	98	4.1	1214	2	US-08-486-273A-54
8	98	4.1	1214	3	US-08-480-474-54
9	98	4.1	1214	3	US-08-940-086A-54
10	98	4.1	1214	4	US-08-940-035A-54
11	98	4.1	1214	4	US-08-935-105A-54
12	98	4.1	1214	4	US-09-648-797-54

C 13	98	4.1	1219	2	US-08-231-193A-50	Sequence 50, Appl
C 14	98	4.1	1219	2	US-08-486-273A-50	Sequence 50, Appl
C 15	98	4.1	1219	3	US-08-480-474-50	Sequence 50, Appl
C 16	98	4.1	1219	3	US-08-940-086A-50	Sequence 50, Appl
C 17	98	4.1	1219	4	US-08-940-035A-50	Sequence 50, Appl
C 18	98	4.1	1219	4	US-08-935-105A-50	Sequence 50, Appl
C 19	98	4.1	1219	4	US-09-648-797-50	Sequence 50, Appl
C 20	98	4.1	1231	2	US-08-231-193A-48	Sequence 48, Appl
C 21	98	4.1	1231	2	US-08-486-273A-48	Sequence 48, Appl
C 22	98	4.1	1231	3	US-08-480-474-48	Sequence 48, Appl
C 23	98	4.1	1231	3	US-08-940-086A-48	Sequence 48, Appl
C 24	98	4.1	1231	4	US-08-940-035A-48	Sequence 48, Appl
C 25	98	4.1	1231	4	US-08-935-105A-48	Sequence 48, Appl
C 26	98	4.1	1231	4	US-09-648-797-48	Sequence 48, Appl
C 27	98	4.1	1236	2	US-08-231-193A-6	Sequence 6, Appl
C 28	98	4.1	1236	2	US-08-486-273A-6	Sequence 6, Appl
C 29	98	4.1	1236	3	US-08-480-474-6	Sequence 6, Appl
C 30	98	4.1	1236	3	US-08-940-086A-6	Sequence 6, Appl
C 31	98	4.1	1236	4	US-08-940-035A-6	Sequence 6, Appl
C 32	98	4.1	1236	4	US-08-935-105A-6	Sequence 6, Appl
C 33	98	4.1	1236	4	US-09-648-797-6	Sequence 6, Appl
C 34	98	4.1	1239	2	US-08-231-193A-52	Sequence 52, Appl
C 35	98	4.1	1239	2	US-08-486-273A-52	Sequence 52, Appl
C 36	98	4.1	1239	3	US-08-480-474-52	Sequence 52, Appl
C 37	98	4.1	1239	3	US-08-940-086A-52	Sequence 52, Appl
C 38	98	4.1	1239	4	US-08-940-035A-52	Sequence 52, Appl
C 39	98	4.1	1239	4	US-08-935-105A-52	Sequence 52, Appl
C 40	98	4.1	1239	4	US-09-648-797-52	Sequence 52, Appl
C 41	98	4.1	1244	2	US-08-231-193A-46	Sequence 46, Appl
C 42	98	4.1	1244	2	US-08-486-273A-46	Sequence 46, Appl
C 43	98	4.1	1244	3	US-08-480-474-46	Sequence 46, Appl
C 44	98	4.1	1244	3	US-08-940-086A-46	Sequence 46, Appl
C 45	98	4.1	1244	4	US-08-940-035A-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-479-309-2
Sequence 2, Application US/09479309
Patent No. 6110691
GENERAL INFORMATION:
APPLICANT: Wang, Xiaodong
TITLE OF INVENTION: Activators of Caspases
FILE REFERENCE: WTS0630
CURRENT APPLICATION NUMBER: US/09/479,309
CURRENT FILING DATE: 2000-01-06
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 239
TYPE: PRT
ORGANISM: human
US-09-479-309-2

Alignment Scores:

Pred. No.: 1.56e-111
Score: 1201.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 49.10%
DB: 3
Gaps: 0

US-09-939-293-1 (1-1358) x US-09-479-309-2 (1-239)
QY 20 ATGGCGGCTCTGAAGAGTGTGGCTGTGGCGGACGCTTAATCTTCAGGTACAGACAG 79
DB 1 MetAlaAlaLeuYsSerTrpLeuSerArgSerValThrSerPhePheArgTyrArgGln 20
QY 80 TGTGTGTGTGTCTGT 139
DB 21 CysLeuCyValProValValAlaAsnPhelYsArgCySPheserGluLeuIleArg 40

OY		140	CCGGCGCAAAAGCTGACGATTCGGTTTGGAGTAACCGTGTCGGGTCTCATTTGCA	199
Db		41	PTOTPHLSYSHTRVAlnrlleGlYPheIyAlThrLeucysAlaValProIIeAla	60
OY		200	CAGAATAACAGACCCTCATTCCCTTAGTACGAAGCATTTGATGAGAGACAGCTCTTTG	259
Db		61	GlnlysserGIuProlHisserLeuSerSerglunlaIeuMetArGArgAlaValSerLeu	80
OY		320	ACTGAATATCTAAGCGCTGTTTTAACCTTAACTCTCTTTAACGCACAATATAAGTTTA	379
Db		101	ThrIuTrlyThrIylsAlaValAlTyThrLeuthrSerleuYlrArGInlyTrhlrSerleu	1200
OY		380	CTTGGGAAAATGAATTCAGAGAGGAGGAAGATGAAGTGTGGCAGGTGATCAATAGAGCCAGA	439
Db		121	LeuGIlysmetAsnserrGIuGIuAspgluValTrprglInvalIlleIGlyAlaArg	1400
OY		440	GCTAGATGACTTCAAACACCAAGACTACTTGAACTGGAAAACCACTTGATGAACTATT	499
Db		141	AlaGIuMetThrSerLyShISgIngluTyTrleuYlsleuGIuTrhlrTrhrrmetThrAla	1600
OY		500	GTTCGCTTTTCAGAGATGGGACGAGGAGCATATCAATCAACCTGGCGGACATTCAGGCCCTCT	559
Db		161	ValGIlyleuSerGIuMetAlaAlaGIuAlaValAlatyrGIuThrGIylaAspgInAlaser	1800
OY		560	AFAACGCCACAGAAATCACATTCACCTGCTGTAACCTGCAGGTGGAAGAGTGCACACAGCTC	619
Db		181	IleThrAlaArgAsnHIsIleGIuIeuValIlysleuGIuValGIuGIuValIHISgInleu	2000
OY		620	TCCCAGAAAGCAAAACCAAGCTGTGGCAGGAAGCACAGTAAAGAAGCTCCGTCAGAAAAACA	679
Db		201	SerrArglySAalagluTrhlrIlySleuAlaGIuAlaGIuInlleGIuGIuleuArGInlySThr	2200
OY		680	CAGAGGAAGGGGAGGAGCGGGCTGAGTCCGAGCAGGAGGCGTAACTGCGTGAAGAT	736
Db		221	GlnIuGIuGIuGIuGIuAlaArgAlaGIuSerGIuGIuAlaTyrlleuArGIuAsp	239
RESULT 2				
US-08-965-903B-2				
Sequence 2, Application US/08965903B				
Patent No. 6060275				
GENERAL INFORMATION:				
APPLICANT: Hachon, Nit				
APPLICANT: Krasnow, Mark A.				
TITLE OF INVENTION: SPROUT PROTEIN AND CODING				
TITLE OF INVENTION: SEQUENCE				
NUMBER OF SEQUENCES: 20				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Dehlinger & Associates				
STREET: 350 Cambridge Ave., Suite 250				
CITY: Palo Alto				
STATE: CA				
COUNTRY: USA				
ZIP: 94306				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Diskette				
COMPUTER: IBM Compatible				
OPERATING SYSTEM: DOS				
SOFTWARE: PASTEO for Windows Version 2.0				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/08/965,903B				
FILING DATE: 07-NOV-1997				
CLASSIFICATION: 435				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: US 60/030232				
FILING DATE: 07-NOV-1996				
ATTORNEY/AGENT INFORMATION:				
NAME: Petithory, Joanne R				
REGISTRATION NUMBER: 42,995				

1	REFERENCE/DOCKET NUMBER: 8600-0177. 30	
2	TELECOMMUNICATION INFORMATION:	
3	TELEPHONE: 650-324-0880	
4	TELEFAX: 650-324-0960	
5	TELEX:	
6	INFORMATION FOR SEO ID NO: 2:	
7	SEQUENCE CHARACTERISTICS:	
8	LENGTH: 591 amino acids	
9	TYPE: amino acid	
10	STRANDEDNESS: single	
11	TOPOLOGY: linear	
12	MOLECULE TYPE: protein	
13	FRAGMENT TYPE: internal	
14	US-08-965-903B-2	
15	Alignment Scores:	
16	Pred. No.: 0.0432	Length: 591
17	Score: 106.00	Matches: 70
18	Percent Similarity: 34.86%	Conservative: 40
19	Best local Similarity: 22.63%	Mismatches: 115
20	Query Match: 4.33%	Indels: 98
21	DB: 3	Gaps: 16
22	US-09-939-293-1 (1-1358) x US-08-965-903B-2 (1-591)	
23	QY 518 GCACGACAAACCTGGATATTCAAATCGGCGGACGAGATCAGCGCTGTATTAACCGCCGAGAAATCAG 577	
24	DB 209 AAlaThrArgSerGlnHisProAlaGlyGlnGlnAspAsnGlyGlnThrThrHisHis 228	
25	QY 578 ATTTCAGCTGGTGAATATGTCAGTGGAAGAGTGTCACACAGCTCTCCCGGAAGACGAAAC 637	
26	DB 229 LeuLeuLeuLeuProGlnArgAsnGlnHisLeuHis--LeuGlnGlnHisGlnHis 247	
27	QY 638 AAGCTGGCAGACAGACAGATGAAAGACTCCGTCAGAAACACAGAGAGAGAGGAGAG 697	
28	DB 248 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnHisLeuGln 267	
29	QY 698 CGGCGCTAGTCGGGAGGAGGAGGCTACG--CGTAGGATTCAGGG 742	
30	DB 268 HisGlnGlnGlnGlnHisAlaAlaGlyLeuAlaThrThrGlnAlaHisSerValGly 287	
31	QY 743 CQTGAGCAGACATGCTCTCTCCCACTCA--GTGGGAAA 781	
32	DB 288 SerAspHisThrAspGlyLeuLeuHisSerHisLeuGlnAsnSerThrThrLysProPro 307	
33	QY 782 GCAGGGGCAATGTCACCCCTGGGCTGGATGCTGTGTCACCCGAGAGAGGC 841	
34	DB 308 AlaSerLysGlnProAlaLeuProAlaGlyMetGlyLeu--Gly 322	
35	QY 842 GGCAGGT-----CTGCGCTGGCCAA 862	
36	DB 323 -LeuGlyLeuGlyLeuGlyLeuAsnGlnProIleIleThrLysGlnProThrProAlaThr 342	
37	QY 863 TCAGGCGAGAGCGCTTTGTGAAGTGTGAAGTGCCTGTGGTCTCAGAGCTTGGCTGAGC 922	
38	DB 342 rGlnLysGlnArgMetHisAlaLeu-GlnGlnLeuGlnProGlyGlyAlaGlyAla 362	
39	QY 923 CTGGTCTTGAACCTTGGGAGCTGCA-----CCGTGTAACTTCAACCCCACTGCG- 975	
40	DB 362 snGlyGly--ProLeuValMetAlaGlyAspProSerLeu--LeuAsnProIleValC 380	
41	QY 976 -----TACAGCTGTCTTACCACATTTTCTTACCTGACA 1009	
42	DB 380 ySProAlaGlySGlyArgGlyArgGlyGlnGlnGlnSerProArgProLeuProGln- 399	
43	QY 1010 CCCAAGACATTTGCTTACCTGTGGGTGACAGAGAGAGTCTTTTGTGATGSC----- 1062	
44	DB 400 -----ThiThrValCysAsnLysThrCysLeuGlySerAlaGlnSerV 414	
45	QY 1063 --CTTAAGTTCAGACAACTGTTTAACTGTTTTCAGCTATTTTACGTGTGTAATAATAT 1120	
46	DB 414 allIleSpryAlaIleCysGlyLeuGlyCysAlaAlaGlyAlaLeu-----P 428	


```

QY 925 CAGGTCCAGCGAGGCTGTGACACAGGAGGCACTCACAGCTCACAAAGGCGTCCGCC 866
Db 940 -----ProProsp-----GlyGlyArgAlaAlaLeuValArg-ALAPro 953
QY 865 TGATTGGCCAGGAGGAGAGAGCTCCGCGCTCTCCGTCGACAGACAGTATGCAACCTT 806
Db 953 OGlnProProGlyArg-----ProProthProGlyProPro-----LeuSerAspVa 969
QY 805 GGGCAGGCTGGCATCTGCCCTTGTCCCTCCACGTAGTGGGAG-----ACAGGCGACT 752
Db 969 lSerArgValSerArgArgProAlaArgProAlaArgTrp-ProValArgThrGlyHisC 989
QY 751 GTGCTCAGGCGCTGCAT----- 735
Db 989 yGGLyArgHisIleuSerAlaSerGluArgProLeuSerProAlaArgCysHisTyrSers 1009
QY 734 -----CCTCAGCAGGTAGAGCTCTCTGTCGACACTCAGCCCGCTCCCTCTCTCTG 680
Db 1009 eRPheroArgAlaAspArgSerGlyArgProPhoLeuProLeuPheProGluProProG 1029
QY 679 TGTTCCTGACGAGCTCTCTA-----TCTGTGCTTC 647
Db 1029 luleuGluAspLeuProLeuGlyProGluGlnLeuAlaArgArgGluAlaLeuLeuA 1049
QY 646 TGCAGCTTGCTGCTTCGCTTCGCGGAGAGCTGGTGCACCTCTCCACCTGCAGTTCA- 588
Db 1049 snAlaAlaTrpAlaArgGlySerArgProSerHisAlaSerLeuProSerValAlaG 1069
QY 587 -----CCAGCTGAATGTATTCCTGCGGTTATAGAGC 554
Db 1069 lAlaAlaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArg 1089
QY 553 CTGAT-----CTCGCCAGTTTATATGACGCTTCTGCTCCAT 515
Db 1089 roAspArgLysHisSerAlaCysArgArgLeuAlaGln--AlaGlnSerMetCysLeuProI 1108
QY 514 CTTCTGAAGAAGCAGCACTGATCCAGTGGTTTCCAGCTTCAT 468
Db 1108 lETyArgGluAlaCysGlnGluGlyGlnAlaGlyAlaProAlaTrpGlnHisArgG 1128
QY 467 --ACTCTTGCTGTTTGAAGTCACTCAGCTGCTGCTCTATGATGACCTGCACACTTC 410
Db 1128 lnhIstValCysLeuHisAlaHisAlaHisLeuProLeuCysTrpGlyAlaValCysProH 1148
QY 409 ATCTCTCTCTCTGAATCA 390
Db 1148 lSLeuProProCysAspSer 1154

RESULT 8
US-08-480-474-54
; Sequence 54, Application US/08480474
; Patent No. 6033865
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA ENCODING
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,474
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9382B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-474-54

Alignment Scores:
Pred. No.: 0.375 Length: 1214
Score: 98.00 Matches: 65
Percent Similarity: 33.58% Conservative: 25
Best Local Similarity: 24.25% Mismatches: 92
Query Match: 4.06% Indels: 86
DB: 3 Gaps: 12

US-09-939-293-1 (1-1358) x US-08-480-474-54 (1-1214)
QY 1009 TGTGAGCTGAATAAATGGGTAAAGCAGCTGT----- 977
Db 913 CysProthProArgSerGlyProSerProCysLeuProthProAspProProProGlu 932
QY 976 -----ACAAGTGGGCTGAATGTTAAACAGGTCAGTCCCAAGGCTAAGAAC 926
Db 933 ProSerProthClyTrpGly----- 939
QY 925 CAGGTCCAGCGAGGCTGTGACACAGAGGCACTCACAGCTCACAAAGGCGTCCGCC 866
Db 940 -----ProProsp-----GlyGlyArgAlaAlaLeuValArg-ALAPro 953
QY 865 TGATTGGCCAGGAGGAGAGAGCTCCGCGCTCTCCGTCGACAGACAGTATGCAACCTT 806
Db 953 OGlnProProGlyArg-----ProProthProGlyProPro-----LeuSerAspVa 969
QY 805 GGGCAGGCTGGCATCTGCCCTTGTCCCTCCACGTAGTGGGAG-----ACAGGCGACT 752
Db 969 lSerArgValSerArgArgProAlaArgProAlaArgTrp-ProValArgThrGlyHisC 989
QY 751 GTGCTCAGGCGCTGCAT----- 735
Db 989 yGGLyArgHisIleuSerAlaSerGluArgProLeuSerProAlaArgCysHisTyrSers 1009
QY 734 -----CCTCAGCAGGTAGAGCTCTCTGTCGACACTCAGCCCGCTCCCTCTCTG 680
Db 1009 eRPheroArgAlaAspArgSerGlyArgProPhoLeuProLeuPheProGluProProG 1029
QY 679 TGTTCCTGACGAGCTCTCTA-----TCTGTGCTTC 647
Db 1029 luleuGluAspLeuProLeuGlyProGluGlnLeuAlaArgArgGluAlaLeuLeuA 1049
QY 646 TGCAGCTTGCTGCTTCGCTTCGCGGAGAGCTGGTGCACCTCTCCACCTGCAGTTCA- 588
Db 1049 snAlaAlaTrpAlaArgGlySerArgProSerHisAlaSerLeuProSerValAlaG 1069
QY 587 -----CCAGCTGAATGTATTCCTGCGGTTATAGAGC 554
Db 1069 lAlaAlaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArg 1089
QY 553 CTGAT-----CTCGCCAGTTTATATGACGCTTCTGCTCCAT 515
Db 1089 roAspArgLysHisSerAlaCysArgArgLeuAlaGln--AlaGlnSerMetCysLeuProI 1108

```

```

Yy      514 CTCTAAGACCAACTGGCGTCATGCACGTGGTTCCAGCTTCCAAGT----- 468
Db      1108 LETTAGTGGLUAlAcysgtngLUGlGlnAlAlaclyAlalatrGrLnInAtgG 1128
               ::::: |||
Yy      467 --ACTTGTGTGTGTTTGAAGTCATCTCAGCTGTGGCTCCATGATGCACSTGCCACATTC 410
               ::::: |||
Db      1128 InHIsValcysLeuHnIstAlnAlnHIsLseuProleucysTrgrLyAlaValcysPrOn 1148
               ::::: |||
Yy      409 ATCTTCCTCTCTGAATTCA 390
               |||||
Db      1148 ISteuPrOProcysAsPser 1154
               |||||

RESULT 9
US-08-940-086A-54
Sequence 54, Application US/08940086A
Patent No. 6111091
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Ellis, Steven B.
APPLICANT: Liaw, Chen W.
APPLICANT: Lu, Chin-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,086A
FILING DATE: 29-SEPT-97
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9383C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 450-8499
INFORMATION FOR SEQ ID NO.: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 1214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-086A-54

Alignment Scores:
Pred. No.:          0.375           Length:       1214
Score:              98.00           Matches:        65
Percent Similarity: 33.58%         Conservative:   25
Best Local Similarity: 24.25%     Mismatches:    92
Query Match:        4.06%          Indels:         86
DB:                 3             Gaps:           12

US-09-939-293-1 (1-1358) x US-08-940-086A-54 (1-1214)
Yy      1009 TGTGTGAGTAATAAAATGGTGTAGAGCACGCTGT----- 977
               |||
               ::: ||| ||| |||

```

```

Db 913 CysProThrProAlaSerGlyProSerProCysLeuProThrProAspProProGlu 932
OY 976 -----ACAGAGTGGGGTGAATAATGTTAAACAGGGTCGACGTGCCAAGGCTAAAGAAC 926
Db 933 ProSerProThrGlyTyrLeu----- 939
OY 925 CAGGTCCAGGCGGACGCGTGAACCCACAGGAGGCGACACAGGTCACACAAAGGCGTCGCGC 866
Db 940 -----ProProAsp-----GlyGlyATgAlaAlaLeuValAlaArg -AlaArg -AlaArg 933
OY 865 TGATTGGCCAGGCGAGGACAGCTCCGCGCTCTTCCGTGACAGACAGTCAAGCTCAACCCCT 806
Db 953 GGLnProProGlyArg-----ProProThrProGlyProPro-----LeuSerAspVal 969
OY 805 GGGCAGGCGTGCATCTGCCCTCTGCTTTCCCACTAGTGGGAG-----ACAGGCGACT 752
Db 969 IserATgValSerATgArgProAlaArgAlaArgTyr -ProValArgThrGlyHisC 989
OY 751 GTCGTCAGGCGCTCAAT----- 725
Db 989 ySGlyATgHisLeuSerAlaSerGlyAlaArgProLeuSerProAlaArgCysHisTyrSer 1009
OY 734 -----CTTCACACAGGTAGGCGCTCTGTCGACACTACAGCGCGCTCCCTCTCTCTG 680
Db 1009 eRProProAlaArgAlaAspArgSerGlyArgProPhoLeuProLeuPhoGluProProG 1029
OY 679 TGTTTTTCAGCGAGCTCTTA-----TCTGTGCTTC 647
Db 1029 LuLeuGluAlaSerProLeuGluArgGlyProGluGluAlaArgAlaArgAlaLeuLeu 1049
OY 646 TGGCAGGCTGGTTCGCTTCCGGGAGAGCTGAGTGCACCTCTTCACACGCGATTCA- 588
Db 1049 smaAlaATrAlaArgGlySerATgProSerHisAlaSerLeuProSerSerValAlaG 1069.
OY 587 -----CCAGCTGAATGTGATTCCTGCGGCTTATAGAGCG 554
Db 1069 LuAlaPhaAlaArgProSerSerLeuProAlaGlyArgHisGlyProAlaCysAlaArg 1089
OY 553 CTGAT-----CTGCGCGAGTTTATATGACAGCTTCTGCTCCAT 515
Db 1089 roAspGlyHisSerAlaCysArgATgLeuAlaGln--AlaGlnSerMetCysLeuPro 1108
OY 514 CTCGTAAAGAACCACTGCAGTCATCCAACTGGTTCCACCTTCAAGT----- 468
Db 1108 LeTgATgAGluAlaCysGlnGluGlyGluAlaAlaGlyAlaProAlaTArgLInHisArg 1128
OY 467 --ACTCTGGTGGTTTGAAGTCATCTCAGCTCGCTGCTCTATGATCAGCTGCACACTTC 410
Db 1128 LInHisValCysLeuHisAlaHisAlaHisLeuProLeuGlySTrGlyAlaValCysProH 1148
OY 409 ATCTCTCTCTCTGAATTC 390
Db 1148 IsLeuProProCysAspSer 1154

RESULT 10
US-08-940-035A-54
: Sequence 54, Application US/08940035A
: Patent No. 6316611
: GENERAL INFORMATION:
: APPLICANT: Daggett, Lorie P.
: APPLICANT: Ellis, Steven B.
: APPLICANT: Law, Chen W.
: APPLICANT: Lu, Chin-Chun
: TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
: FULL OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
: NUMBER OF SEQUENCES: 63
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Heller Ehrman White & Mcauliffe
: STREET: 4250 Executive Square, 7th Floor
: CITY: La Jolla
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92037

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,035A
FILING DATE: 29-SEPT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 1214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-035A-54

Alignment Scores:
Pred. No.: 0.375 Length: 1214
Score: 98.00 Matches: 65
Percent Similarity: 33.588 Conservative: 25
Best Local Similarity: 24.258 Mismatches: 92
Query Match: 4.064 Indels: 86
Gaps: 12
DB:

US-09-939-293-1 (1-1358) x US-08-940-035A-54 (1-1214)

QY 1009 TGTGAGGTAAAAATGGGTAGACGACGTGT----- 977
DB 913 CysProthPrProArGserGlyProSerProCysLeuProThProAsrProProProGlu 932
QY 976 -----ACAGAGTGGGGTGAATGTAAACAGGGTGCACATGCCCAAGGGCTAAGAC 926
DB 933 ProSerProThGlyTrpIly----- 939
QY 925 CAGGTCCAGCGCAGCAGCTGACACAGAGGAGGACACTACAGCTCACAAGGCGTCCGCC 866
DB 940 -----ProProAsp-----GlyGlyArgAlaAlaLeuValArgArg-AlaPr 953
QY 865 TGAATGGCCAGGAGCAGGACCTGCGGCTCTTCTCGGTGCACAGACAGTCCCAACCT 806
DB 953 OGInProProGlyArg-----ProProThProGlyProPro-----LeuSerAspVa 969
QY 805 GGGCAGGCGGAGTGTGCGCCCTGCTTCCCACTAGTGGGAG-----AAGGGCAAT 752
DB 966 lSerArgValSerArgArgProAlaTrpGluAlaArgTrp-ProValArgThnGlyHnSc 989
QY 751 GTGCTCAGGCGCTCAAT----- 735
DB 989 ysgLlyArgHnSerAlaSerGluArgProLeuSerProAlaArgCysHnIySerS 1009
QY 734 -----CCTCAGCCAGGAGGCTCTGCTCCAGCTCAGCGCCCTCCCTCTCTCTG 680
DB 1009 eRPhetrOrArgAlaAspArgSerGlyArgProPhelAuProLeuPhelProGluProProG 1029
QY 679 TGTTTTCGACGAGCTCTTCA-----TCTGTCCTTC 647
DB 1029 luleuGluAspLeuProLeuLeuGlyProGluGluLeuAlaArgArgGluAlaLeuLeu 1049
QY 646 TCCAGACTTGTCTTGTCTTCCGGAGAGCTGTGTCACCTCTTCCACCTGCAGTTTCA- 588

DB 1049 snAlaAlaTrpAlaArgGlySerArgProSerHnAlaSerLeuProSerSerValAlag 1069
QY 587 -----CAGCTGAATGTATTCCTGGCGGTTATAGAGCC 554
DB 1069 lAlaRPhelAlaArgProSerSerLeuProAlaGlyCysThrclProAlaCysAlaArgP 1089
QY 553 CTGAT-----CTGGCCAGTTTATATMGACGCTTCTGTCGCCAT 515
DB 1089 roAspGlyHnSerAlaCysArgArgLeuAlaGln---AlaGlnSerMetCysLeuProI 1108
QY 514 CTGTAAAGACCAACTGCATCATCCAAAGTGGTTCCACCTTCAAGT----- 468
DB 1108 letyArgGluAlaCysGlnGluGluAlaGlyAlaRProAlaTrpGlnHnIArgS 1128
QY 467 --ACTCTTGCTGTGTAAGTATCTCAGCTGTGCTGCTTAATGATCAGCTGCCACACTTC 410
DB 1128 lnhIsvAlCysLeuHnIsvAlaHnIsvAlaHnIsvLeuProLeuCysTrpGlyAlaValCysProH 1148
QY 409 ATCTTCCTCTGTAATCA 390
DB 1148 lsvLeuProCysAspSer 1154

RESULT 11
US-08-935-105A-54
Sequence 54, Application US/08935105A
Patent No. 637660
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESS: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,105A
FILING DATE: 29-SEPT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 1214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-935-105A-54
Alignment Scores:
Pred. No.: 0.375 Length: 1214
Score: 98.00 Matches: 65

Percent Similarity: 33.58% Conservative: 25
 Best Local Similarity: 24.25% Mismatches: 92
 Query Match: 4.06% Indels: 86
 DB: 4 Gaps: 12

US-09-939-293-1 (1-1358) x US-08-935-105A-54 (1-1214)

```

QY 1009 TGTGAGGTAAATAATGGTAAGACAGCTGT----- 977
    |||  |||  |||  |||
Db 913 CysProthPrProAArgSerGlyProSerProCysLeuProthPrProAArgProProGlu 932
QY 976 -----ACAGAGTGGGGTGAATGTAAACAGGGTGACGTGCCAAGGGCTAAGAAC 926
    |||  |||  |||  |||
Db 933 ProSerProthPrGlyTrpGly----- 939
QY 925 CAGGTCCAGGCCAAGCTGAGAACACAGAGAGACACTCACAGCTCAACAAAGGCTGTGCC 866
    |||  |||  |||  |||
Db 940 -----ProProAArg-----GlyGlyArgAlaAlaLeuValArgArg-AlaPr 953
QY 865 TGATTTGGCAGGAGACAGCTGCGGCTCTTCGCTGCACAGACAGTCCCAACCT 806
    |||  |||  |||  |||
Db 953 OGlnProProGlyArg-----ProProthPrProGlyProPro-----LeuSerAspVa 969
QY 805 GGGCAGGGTGGCATCTGCCCCCTGCTTCCCACTGAGTGGGAG-----ACAGGGCAGT 752
    |||  |||  |||  |||
Db 969 lSerArgValSerArgArgProAlaTrpGluAlaArgTrp-ProValArgThrGlyHisc 989
QY 751 GTGCTCAGGCCCCCAAT----- 735
    |||  |||  |||  |||
Db 989 ysglyArgHiscLeuSerAlaSerGluArgProLeuSerProAlaArgCysHiscTyrSers 1009
QY 734 -----CCACAGCAGGTAGGGCTCTGCTCCGACTCAGCCGCGCTCCCTCTCTCTG 680
    |||  |||  |||  |||
Db 1009 erPrProAlaArgAlaSerArgSerGlyArgProHiscLeuProLeuPrProGluProG 1029
QY 679 TGTTTTTCGACGAGAGCTCTCTTA-----TCTGTGCTTC 647
    |||  |||  |||  |||
Db 1029 luleGluAspLeuProLeuLeuGlyProGluGlnLeuAlaArgArgGluAlaLeuLeuA 1049
QY 646 TGCCAGCTGTGCTCTCTCTCCGAGAGAGCTGGTGCACCTCTCCACCTCACTTTCA 588
    |||  |||  |||  |||
Db 1049 snAlaAlaTrpAlaArgGlySerArgProSerHiscAlaSerLeuProSerSerValAlaG 1069
QY 587 -----CCAGCTGAATGTGATTCGCGGGTTATAGAGGC 554
    |||  |||  |||  |||
Db 1069 lualaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArgP 1089
QY 553 CTGAT-----CTGCGCAGTTTGAATAGCAGCTTCTGCTGCAT 515
    |||  |||  |||  |||
Db 1089 roAspGlyHiscSerAlaCysArgArgLeuAlaGln--AlaGlnSerMetCysLeuProI 1108
QY 514 CTCGTAAGACCACTGACAGTCAATCCAGTGTTCACGCTCAAGT----- 468
    |||  |||  |||  |||
Db 1108 leryArgGluAlaCysGlnGluGlnAlaGlyAlaProAlaTrpGlnHiscArgG 1128
QY 467 --ACCTGTGGTGTGTAAGATCACTCAGCTGTGCTGATGATCAGCTCCACACATTC 410
    |||  |||  |||  |||
Db 1128 lnhIsvAlCysLeuHiscAlaHiscAlaHiscLeuProLeuCysTrpGlyAlaValCysProH 1148
QY 409 ATCTTCTCTCTCTCAATCA 390
    |||  |||  |||  |||
Db 1148 lslEuProProCysAspSer 1154

```

RESULT 12

US-09-648-797-54

; Sequence 54, Application US/09648797

; Patent NO. 6469142

; GENERAL INFORMATION:

APPLICANT: Daggett, Lorrie P.

Ellis, Steven B.

Liaw, Chen W.

Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCulliffe

STREET: 4250 Executive Square, 7th Floor

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/648,797

FILING DATE: 28-Aug-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/940,086A

FILING DATE: 29-SEPT-97

APPLICATION NUMBER: US 08/231,193

FILING DATE: 20-APR-1994

APPLICATION NUMBER: US 08/052,449

FILING DATE: 20-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 24735-9383C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 450-8409

TELEFAX: (619) 450-8499

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:

LENGTH: 1214 amino acids

TYPE: amino acid

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 54:

US-09-648-797-54

US-09-939-293-1 (1-1358) x US-09-648-797-54 (1-1214)

QY 1009 TGTGAGGTAAATAATGGTAAGACAGCTGT----- 977

Db 913 CysProthPrProAArgSerGlyProSerProCysLeuProthPrProAArgProProGlu 932

QY 976 -----ACAGAGTGGGGTGAATGTAAACAGGGTGACGTGCCAAGGGCTAAGAAC 926

Db 933 ProSerProthPrGlyTrpGly----- 939

QY 925 CAGGTCCAGGCCAAGCTGAGAACACAGAGAGACACTCACAGCTCAACAAAGGCTGTGCC 866

Db 940 -----ProProAArg-----GlyGlyArgAlaAlaLeuValArgArg-AlaPr 953

QY 865 TGATTTGGCAGGAGACAGCTGCGGCTCTTCGCTGCACAGACAGTCCCAACCT 806

Db 953 OGlnProProGlyArg-----ProProthPrProGlyProPro-----LeuSerAspVa 969

QY 805 GGGCAGGGTGGCATCTGCCCCCTGCTTCCCACTGAGTGGGAG-----ACAGGGCAGT 752

Db 969 lSerArgValSerArgArgProAlaTrpGluAlaArgTrp-ProValArgThrGlyHisc 989

QY 751 GTGCTCAGGCCCCCAAT----- 735

Db 989 ysglyArgHiscLeuSerAlaSerGluArgProLeuSerProAlaArgCysHiscTyrSers 1009

Alignment Scores:

Pred. No.: 0.375

Score: 98.00

Percent Similarity: 33.58%

Best Local Similarity: 24.25%

Query Match: 4.06%

DB: 4

Gaps: 12

Length: 1214

Matches: 65

Conservative: 25

Mismatches: 86

Indels: 12

[illegible]

RESULT 13

```

US-08-231-193A--50
: Sequence 50, Application US/08231193A
: Patent No. 5849895
:
: GENERAL INFORMATION:
:
: APPLICANT: Daggett, Lorrie P.
: APPLICANT: Ellis, Steven B.
: APPLICANT: Liaw, Chen W.
: APPLICANT: Lu, Chin-Chun
:
: TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
: TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
:
: NUMBER OF SEQUENCES: 63
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: CA
:
: COUNTRY: U.S.A.
: ZIP: 92101-2926
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/231,193A
: FILING DATE: 20-APR-1994
: CLASSIFICATION: 536
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/052,459
: FILING DATE: 20-APR-1993
: CLASSIFICATION: 536
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-9383
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-238-0999
: TELEFAX: 619-238-0062
:
: INFORMATION FOR SEQ ID NO: 50:

```

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 1219 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-231-193A-50

```

Alignment Scores:

Pred. No.:	0.376	Length:	1219
Score:	98.00	Matches:	65
Percent Similarity:	33.58%	Conservative:	25
Best local Similarity:	24.25%	Mismatches:	92
Query Match:	4.06%	Indels:	86
DB:	2	Gaps:	12

US-09-939-293-1 (1-1358) x US-08-231-193A-50 (1-1219)

QY 1009 TGTGAGGTAAAAAATGGGTAAAGACGACTGT----- 977
::: ||| |||

Db 918 CysProThrProArgSerGlyProSerProCysLeuProThrProAspProProGlu 937

QY 976 -----ACAGAGTGGGGTGAATGTTAAACAGGGTGCAGTGCACCAAGGGCTAAGAAC 926

```

Db      938 ProSerProThrGlyTyrGly----- 944
      ||| |||||

```

925 CAGGTCCAGCGCAAGCCTGAGAGCCACAGAGAGGCCTCACAAGTCACAAAGGCGTGGCC 866

[illegible]

```
00000000-PROFIOBSP-----GUYyylghlatataduuvaialngal-g-MIARf 550
```

QY 865 TGATTGGCCAGGCGCAGACCTGCGCGCTCTTTCGGGTGCACAGACATGCCACACCT 806

Db 958 OGI^{Pro}ProGI^{Var}-----Pro^{Pro}Th^{Pro}GI^{Pro}GI^{Pro}Pro-----LeuSerAspVa 974

Q7 805 GGGCAGGGTGGCATCTGCCCTGCTTCCCACTGAGTGGGAG-----ACAGGGCACT 752Z

Db 974 1SerArgValSerArgArgProAlaTrpGluAlaIaArgTrp-ProValaArgThrGluNisC 994

QY 751 GTGCTCAGGCCCTCAAT----- 735

Db 994 VSG[VA]VGHISLeuSerAlaSerG[VA]VqProLeuSerProAlaArGcVshISTVrSers 101

734 600

27 CCACACGAGGATGGGCTCTGCTCCGACTACGCGCGCTCCCTCCCTCCG 680
 ||| ||| ||| ||| |||||

Db 1014 erpHeProArGaLaaspArgSerGlyArgProPheLeuProLeuPheProGluProProG 103

QY 679 TGTTCCTGACGACCTCTTCTA-----TCTGTGCTTC 647
|||||
.....

Db 1034 IuleuGIuAspIeuProIeuLeuGIyProGIuGIInLeuAlaArgIaArgGIuAlaLeu 105

646 TGCCAGCTTGTTTCTGCTTTCGGGAGAGCTGCTGACACCTTTCACCTGCAGTTTCA- 588

Db 1054 snAlaAlaTrpAlaAlaArgGlySerAlaProSerGlnHisAlaSerLeuProSerValAlaG 107

587 -----CCACGCGAATCTGATTCGCGCGGCTTATAGAGCG 554

[illegible]

Db 10/4 IuAlaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArg 109

553 CTGAT-----CTGCCCCAGTTTGATATGACAGCTTCTGCTGCCAT 515

Db 1094 roaspGLHISserAlaCysArgArgLeuAlaIn---AlaGlnserMetCysLeuProI 111

514 CTCTGAAGACCAACTGCAGTCATCCAAAGTGGTTTCAGCTTCAAGT----- 468

```

Db 1113 IETyRArgGLuaIacysGInGluGlyGluGlnAlaGlyAlaProAlaTyrGlnHisArg 1133
      ::::: ||| :::: ||| ::::

```

467 --ACTCTGCTGTTTTGAAGTCATCTCAGCCTGGCTCCTATGATCATCTGCCACACTTC 410

```
Db      1133  [u]cuf[u]m[u]c[u]a[u]n[u]s[u]a[u]l[u]a[u]t[u]e[u]r[u]o[u]d[u]r[u]o[u]c[u]u[s]t[u]r[u]o[u]c[u]i[u]a[u]y[a]y[a]CusBroch 1155
        ::   ::   ::   ||   ||   |
x7      107  .....::.....
```

[illegible]

QY 409 ATCTTCCTCCTGTAATCA 390
|||||
:::|

Db 1153 isLeuProProCysaspSer 1159

```

RESULT 14
US-08-486-273A-50
; Sequence 50, Application US/08486273A
; Patent No. 5985586
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1219 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-273A-50

Alignment Scores:
Pred. No.: 0.376 Length: 1219
Score: 98.00 Matches: 65
Percent Similarity: 33.58% Conservative: 25
Best Local Similarity: 24.25% Mismatches: 92
Query Match: 4.06% Indels: 86
Dbs: 2 Gaps: 12

US-09-939-293-1 (1-1358) x US-08-486-273A-50 (1-1219)
QY 1009 TGTGAGGTAAATAAGGCTAAGAGCAGCTGT----- 977
||| :||| ||| |||
Db 918 CysProThrProArGserGlyProSerProCysLeuProThrProAspProProGlu 937
976 -----ACAGATGGGGGTGAATGTAAACAGGGGTGCAGTCCCAAGGGCTAAGAAC 926
||| ||| |||
Db 938 ProSerProThrGlyTyrGly----- 944
925 CAGTCCAGGCGCAAGCTGAGACACAGAGGACACTCAGACTGCACAAAGGCGTCTGCC 866
||| :||| ||| |||
Db 945 -----ProProAsp-----GlyGlyArgAlaAlaLeuValArgArg-AlaPr 958
||| :||| ||| |||
QY 865 TGATTGCCAGGCGAGGACCGGCTCTTCGTCGGTGACACAGACAGTCAATGCCAACCT 806
| ||||| :||| |||
Db 958 OGlnProProGlyArg-----ProProThrProGlyProPro-----LeuSerAspVa 974

```

```

QY 805 GGCAGGGTGAGCATGTGCCCCCTGTTCCCACTGATGGGAG-----ACAGGCGAGT 752
: ||||| :||| |||
Db 974 lSerArgValSerArgArgProAlaTrpGluAlaArgTrp-ProValArgThrGlyHisC 994
QY 751 GTGCTCAGGCCCTCAAT----- 735
|| :||| :|||
Db 994 ysglyArgHisLeuSerAlaSerGluArgProLeuSerProAlaArgCysHisTyrSers 1014
QY 734 -----CCTCAGCGAGTGAAGCTCTGCTGCTGAGTACAGCCGCGCTCCCTGCTG 680
||| ||| ||| ||| ||| ||| ||| |||
Db 1014 erPheProArgAlaAspArgSerGlyArgProPheLeuProLeuPheProGluProProG 1034
QY 679 TGTTCCTGACGGAGCTCTCTA-----TCGTGCTTC 647
||| ||| |||
Db 1034 luleGluAspLeuProLeuLeuGlyProGluGlnLeuAlaArgArgGluAlaLeuLeuA 1054
QY 646 TGCAGCTGGTTTCTGCTTTCCGGGAGAGCTGTGCACCTTTCACCTGCAGTTTCA- 588
||| ||| :||| ||| ||| ||| |||
Db 1054 snAlaAlaTrpAlaArgGlySerArgProSerHisAlaSerLeuProSerSerValAlaG 1074
QY 587 -----CCAGCTGAATGTGATTCCTGGCGGTTATAGGC 554
||| ||| ||| |||
Db 1074 lAlaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArgP 1094
QY 553 CTGAT-----CTGCGCCAGTTTGATATGCAGCTTCTGCTGCCAT 515
||| ||| ||| |||
Db 1094 roAspGlyHisSerAlaCysArgArgLeuAlaGln---AlaGlnSerMetCysLeuProI 1113
QY 514 CTCTGAAGACCACTGCAGTCAATCAAGTGTTCAGCTTCAAGT----- 468
: :||| :|||
Db 1113 leryArgGluAlaCysGlnGlnGluGlnAlaGlyAlaProAlaTrpGlnHisArgG 1133
QY 467 --ACTCTGGGTGTTTAAGTCATCTCAGCTGCTGCTGCTATGATCACTGCCAGCTTC 410-
: :||| :||| ||| |||
Db 1133 lnhlsvAlCysLeuHlslAlaHlslAlaHlslAlaHlslAlaHlslAlaHlslAlaHlsl 1153
QY 409 ATCTTCCTCCTGATTCATCA 390
||| ||| ||| |||
Db 1153 lslLeuProProCysAspSer 1159

RESULT 15
US-08-480-474-50
; Sequence 50, Application US/08480474
; Patent No. 6033865
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA ENCODING
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,474
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9382B
; TELECOMMUNICATION INFORMATION:

```


TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 1219 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-474-50

Alignment Scores:
Pred. No.: 0.376 Length: 1219
Score: 98.00 Matches: 65
Percent Similarity: 33.58% Conservative: 25
Best Local Similarity: 24.25% Mismatches: 92
Query Match: 4.06% Indels: 86
DB: 3 Gaps: 12

US-09-939-293-1 (1-1358) x US-08-480-474-50 (1-1219)

```
QY 1009 TGTGAGGTAAATAATGGGTAGACAGCTGT----- 977
Db 918 CysProThrProArgSerGlyProSerProCysLeuProThrProAspProProGlu 937
QY 976 -----ACAGAGTGGGGTGAATGTTAAACAGGGTGCAGTGCACAGGGCTAAGAAC 926
Db 938 ProSerProThrGlyTyrGly----- 944
QY 925 CAGGTCCAGGCGACGCTGAGACACAGAGGACCTCACAGCTCACAAAGCGCTCGCC 866
Db 945 -----ProProAsp-----GlyGlyArgAlaAlaLeuValArgArg-AlaArg 958
QY 865 TGAATGGCCAGGCGAGACCTGCCCTCTTCGCGTGCGACAGACAGTATGCCAACCT 806
Db 958 ogInProProGlyArg-----ProProThrProGlyProPro-----LeuSerAspVa 974
QY 805 GGCAGGCGTGGCATTCGCCCTTCCTCCCTCCACTGAGTGGGGAG-----ACAGGCGAGT 752
Db 974 lSerArgValSerArgArgProAlaArgTyr-ProValArgThrGlyHisC 994
QY 751 GTGCTCAGGCGCTCAAT----- 735
Db 994 yscLysArgHisLeuSerAlaSerGluArgProLeuSerProAlaArgCysHisTyrSers 1014
QY 734 -----CCTCAGCGAGGTAGCGCTCTGCTCCGACCTCAGCCCGCTCTCCCTCTCTG 680
Db 1014 erPheProArgAlaAspArgSerGlyArgProPheLeuProLeuPheProGluProProG 1034
QY 679 TGTTTTCTGACGGAGCTCTCTA-----TCTGTGCTTC 647
Db 1034 LuLeuGluAspLeuProLeuLeuGlyProGluGlnLeuAlaArgArgGluAlaLeuLeuAla 1054
QY 646 TGGCAGCTTGCTTCTGCTTCCGGGAGAGCTGTCGACCTCTCCACCTGCACTTCA 588
Db 1054 smaLaIatArgAlaArgGlySerArgProSerHisAlaSerLeuProSerSerValAlaG 1074
QY 587 -----CCAGCTGAATGTGATTCCTGGCGGTTATAGAGC 554
Db 1074 LuAlaPheAlaArgProSerSerLeuProAlaGlyCysThrGlyProAlaCysAlaArgP 1094
QY 553 CTGAT-----CTGGCCAGTTTGATATGACAGCTCTGCTGCCAT 515
Db 1094 roAspGlnHisSerAlaCysArgArgLeuAlaGln--AlaGlnSerMetCysLeuProI 1113
QY 514 CTCGTGAAGACCACTGCACATCTCCAACTGCTTCCAGCTTCACAT----- 468
Db 1113 letuArgArgGluAlaCysGlnGluGlyGlnAlaGlyAlaArgAlaArgGlnHisArgG 1133
QY 467 --ACTCTTGCTGTTTGAAGTATCTCAGCTGCTGGCTCTATGATGACACTGCCACACTTC 410
Db 1133 lnhIsvAlCysLeuHisAlaHisAlaHisLeuProLeuCysTyrGlyAlaValCysProH 1153
QY 409 ATCTCTCTCTCTCTGAATTCA 390
```

Db 1153 lshLeuProProCysAspSer 1159

Search completed: February 20, 2003, 15:03:39
Job time : 36.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 20, 2003, 15:01:24 ; Search time 16.5 Seconds

(without alignments)
4205.500 Million cell updates/sec

Title: US-09-939-293-1

Perfect score: 2446

Sequence: 1 ggcgtccgcgcgtcgcacaa.....ggacttaacacagaanaaaa 1358

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 280518

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09939293/runat_20022003_111513_15000/app_query.fasta.1.1543
-DB=Published_Applications_AA -QFM=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09939293.eccn1_1.11 -runat_20022003_111513_15000
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	46.5	227	10	US-09-925-297-591
2	1016	41.5	237	10	US-09-798-116-2
3	1016	41.5	237	10	US-09-798-116-4
4	1004	41.0	202	10	US-09-798-116-7

5	869	35.5	177	10	US-09-798-116-8	Sequence 8, Appl
6	773	31.6	177	10	US-09-798-116-6	Sequence 9, Appl
7	333	13.6	84	10	US-09-798-116-9	Sequence 2, Appl
8	190	7.8	40	10	US-09-939-293-2	Sequence 8, Appl
9	186	7.6	39	10	US-09-939-293-8	Sequence 11, Appl
10	164	6.7	35	10	US-09-939-293-11	Sequence 24, Appl
11	157	6.4	32	10	US-09-798-116-24	Sequence 10, Appl
12	148.5	6.1	73	10	US-09-798-116-10	Sequence 7, Appl
13	141	5.8	30	10	US-09-939-293-7	Sequence 25, Appl
14	137	5.6	29	10	US-09-798-116-25	Sequence 701, App
15	111	4.5	401	9	US-09-764-868-701	Sequence 17, App
16	107	4.4	1509	10	US-09-901-940-2	Sequence 54, Appl
17	103.5	4.2	317	10	US-09-789-561-171	Sequence 54, Appl
18	98	4.1	1214	9	US-09-945-901-54	Sequence 54, Appl
19	98	4.1	1214	9	US-10-007-747-54	Sequence 54, Appl
20	98	4.1	1214	9	US-10-038-937-54	Sequence 50, Appl
21	98	4.1	1219	9	US-09-945-901-50	Sequence 50, Appl
22	98	4.1	1219	9	US-10-007-747-50	Sequence 48, Appl
23	98	4.1	1219	9	US-10-038-937-50	Sequence 48, Appl
24	98	4.1	1231	9	US-09-945-901-48	Sequence 48, Appl
25	98	4.1	1231	9	US-10-007-747-48	Sequence 48, Appl
26	98	4.1	1231	9	US-10-038-937-48	Sequence 6, Appl
27	98	4.1	1236	9	US-09-945-901-6	Sequence 6, Appl
28	98	4.1	1236	9	US-10-007-747-6	Sequence 6, Appl
29	98	4.1	1236	9	US-10-038-937-6	Sequence 52, Appl
30	98	4.1	1239	9	US-09-945-901-52	Sequence 52, Appl
31	98	4.1	1239	9	US-10-007-747-52	Sequence 52, Appl
32	98	4.1	1239	9	US-10-038-937-52	Sequence 46, Appl
33	98	4.1	1244	9	US-09-945-901-46	Sequence 46, Appl
34	98	4.1	1244	9	US-10-007-747-46	Sequence 46, Appl
35	98	4.1	1244	9	US-10-038-937-46	Sequence 2, Appl
36	94.5	3.9	503	9	US-10-078-547-2	Sequence 193, App
37	94	3.8	218	10	US-09-789-561-193	Sequence 6, Appl
38	93.5	3.9	738	10	US-09-978-979-6	Sequence 700, App
39	93.5	3.8	2139	10	US-09-727-384-6	Sequence 7, Appl
40	93	3.9	200	10	US-09-925-297-700	Sequence 47, Appl
41	93	3.8	300	10	US-09-919-603-7	Sequence 14, Appl
42	92.5	3.8	826	9	US-10-084-988-47	Sequence 16, Appl
43	92	3.8	663	9	US-10-080-960-14	Sequence 2, Appl
44	92	3.8	1617	10	US-09-784-358-16	
45	92	3.8	1691	10	US-09-784-358-2	

ALIGNMENTS

RESULT 1
US-09-925-297-591
; Sequence 591, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 591
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-591
Alignment Scores: 1.26e-92 Length: 227


```

: CURRENT APPLICATION NUMBER: US/09/798.116
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: AU P05995/00
: PRIOR FILING DATE: 2000-03-02
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 9
: LENGTH: 84
: TYPE: PRT
: ORGANISM: Rattus sp.
US-09-798-116-9

Alignment Scores:
Pred. No.: 1.08e-21 Length: 84
Score: 333.00 Matches: 68
Percent Similarity: 86.05% Conservative: 6
Best Local Similarity: 79.07% Mismatches: 10
Query Match: 13.61% Indels: 2
DB: 10 Gaps: 1

US-09-939-293-1 (1-1358) x US-09-798-116-9 (1-84)
QY 20 ATGCGCGCTGAGAGAGTGGCTGTGCGCAGCGTAACCTTATCTCAGGTACAGACAG 79
Db 1 MetAlaAlaLeuArgSerThrPheThrValThrPheLeuPheArgTyrGlyGln 20
QY 80 TGTGTGTGTCTCTGCTGTGGCTTAAGAACGCGTGTCTCAGATTTGATAGA 139
Db 21 ArgPhe-----ProValSerAlaSerLysLysArgCysPheSerGluLeuLeuArg 38
QY 140 CCATGGCACAACAGTGTGACGATTTGGAGTAAACCGTGTGCGCTTCCATTGCA 199
Db 39 ProThrPheLysThrMetLeuThrGlyValThrLeuGlyAlaValProIleAla 58
QY 200 CAGAAATCAGAGCTCATTCCCTTAGTAGTGAAGCATTGATGAGAGAGACAGTGTCTTG 259
Db 59 GlnLysSerGluProGlnSerLeuSerAsnGlnAlaLeuMetArgAlaValSerLeu 78
QY 260 GTAACAGATACAGCTCT 277
Db 79 ValThrAsnSerThrSer 84

RESULT 8
US-09-939-293-2
: Sequence 2, Application US/09939293
: Patent No. US20020132786A1
: GENERAL INFORMATION:
: APPLICANT: Alnemri, Emed S.
: TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
: FILE REFERENCE: 480140.465
: CURRENT APPLICATION NUMBER: US/09/939.293
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-939-293-2

Alignment Scores:
Pred. No.: 3.53e-09 Length: 40
Score: 190.00 Matches: 40
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.77% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1 (1-1358) x US-09-939-293-2 (1-40)
QY 185 GCGGTTCTATTGACAGAAATCAGAGCGCTCATTCCTTAGTAGTGAAGCATTGATGAGG 244
|||||
```

```

Db 1 AlaValProIleAlaGlnLysSerGluProHisSerLeuSerGlnAlaLeuMetArg 20
QY 245 AGACGAGTGTCTTTGGTAACAGATAGCACCTCTACCTTTCTCTCAGACCATATGCG 304
Db 21 ArgAlaValSerLeuValThrAspSerThrSerThrPheLeuSerGlnThrThrTyrAla 40

RESULT 9
US-09-939-293-8
: Sequence 8, Application US/09939293
: Patent No. US20020132786A1
: GENERAL INFORMATION:
: APPLICANT: Alnemri, Emed S.
: TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
: FILE REFERENCE: 480140.465
: CURRENT APPLICATION NUMBER: US/09/939.293
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 39
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-939-293-8

Alignment Scores:
Pred. No.: 7.9e-09 Length: 39
Score: 186.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.60% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1 (1-1358) x US-09-939-293-8 (1-39)
QY 185 GCGGTTCTATTGACAGAAATCAGAGCGCTCATTCCTTAGTAGTGAAGCATTGATGAGG 244
Db 1 AlaValProIleAlaGlnLysSerGluProHisSerLeuSerGlnAlaLeuMetArg 20
QY 245 AGACGAGTGTCTTTGGTAACAGATAGCACCTCTACCTTTCTCTCAGACCATAT 301
Db 21 ArgAlaValSerLeuValThrAspSerThrSerThrPheLeuSerGlnThrThrTyr 39

RESULT 10
US-09-939-293-11
: Sequence 11, Application US/09939293
: Patent No. US20020132786A1
: GENERAL INFORMATION:
: APPLICANT: Alnemri, Emed S.
: TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
: FILE REFERENCE: 480140.465
: CURRENT APPLICATION NUMBER: US/09/939.293
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 35
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-939-293-11

Alignment Scores:
Pred. No.: 6.67e-07 Length: 35
Score: 164.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.70% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1 (1-1358) x US-09-939-293-11 (1-35)
QY 185 GCGGTTCTATTGACAGAAATCAGAGCGCTCATTCCTTAGTAGTGAAGCATTGATGAGG 244
|||||
```

```
Db 1 AlavAlproilLeaGlnlySserGluProHisSerLeuSerSerGluAlaLeuMetArg 20
Qy 245 AGACAGCTGCTTGTGTAACAGATAGACACCTCTACCTTCTCTCT 289
Db 21 ArgAlaValSerLeuValThrAspSerThrSerThrPheLeuSer 35
RESULT 11
US-09-798-116-24
; Sequence 24, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekerl, Paul
; APPLICANT: Vaux, David
; TITLE OF INVENTION: No. US20020110851A1 Polypeptides, Modulatory Agents Therefor ar
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU PQ5995/00
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 32
; TYPE: PRT
; ORGANISM: synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: M is methionine sulfoxide
; LOCATION: (14)..(14)
; OTHER INFORMATION: M is methionine sulfoxide
US-09-798-116-24
Alignment Scores:
Pred. No.: 2.68e-06 Length: 32
Score: 157.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.42% Indels: 0
DB: 10 Gaps: 0
US-09-939-293-1 (1-1358) x US-09-798-116-24 (1-32)
Qy 476 CTGGAACCACTGTGAGCTGCACTGCTTCAGAGATGAGCAGACGATGCATAT 535
Db 1 LeuGlnThrThrTrpMetThrAlaValGlyLeuSerGluMetAlaAlaGluAlaLayr 20
Qy 536 CAACGTGGCGAGATCAGGCGCTCTATACCGCCAGG 571
Db 21 GlnThrGlyAlaAspGlnAlaSerIleThrAlaArg 32
RESULT 12
US-09-798-116-10
; Sequence 10, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekerl, Paul
; APPLICANT: Vaux, David
; TITLE OF INVENTION: No. US20020110851A1 Polypeptides, Modulatory Agents Therefor ar
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU PQ5995/00
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 73
; TYPE: PRT
```

```
; ORGANISM: Platicthys flesus
US-09-798-116-10
Alignment Scores:
Pred. No.: 2.02e-05 Length: 73
Score: 148.50 Matches: 30
Percent Similarity: 81.97% Conservative: 20
Best Local Similarity: 49.18% Mismatches: 8
Query Match: 6.07% Indels: 3
DB: 10 Gaps: 1
US-09-939-293-1 (1-1358) x US-09-798-116-10 (1-73)
Qy 173 GTAACCGTGTCGGCTTCATATGACAG-----AAATCAGACCTCATTCCTT 223
Db 13 MetSerIleAlaSerLeuSerValAlaArgGlyLeuPheThrGlnValGluThrLeu 32
Qy 224 AGTAGTCAGCATTTGATGAGAGAGACAGTGTCTTGTGTAACATATACCTCTT 283
Db 33 ThrHisAspSerLeuIleArgAlaValSerValValThrAspSerSerThrPhe 52
Qy 284 CTCCTCAGACCATATATGCTGATGAGCTATTAATCTGAATATTAAGCGTGTAT 343
Db 53 LeuSerGlnThrThrLeuAlaLeuIleAspAlaLeuThrAspTyrSerLysAlaValHis 72
Qy 344 ACC 346
Db 73 Thr 73
RESULT 13
US-09-939-293-7
; Sequence 7, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emed S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-7
Alignment Scores:
Pred. No.: 6.78e-05 Length: 30
Score: 141.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.76% Indels: 0
DB: 10 Gaps: 0
US-09-939-293-1 (1-1358) x US-09-939-293-7 (1-30)
Qy 185 GCGGTCTCTATTGACAGAAATCAGACGCTCATTCCTTGTAGTGAAGATTGAGG 244
Db 1 AlavAlproilLeaGlnlySserGluProHisSerLeuSerSerGluAlaLeuMetArg 20
Qy 245 AGACAGCTGCTTGTGTAACAGATAGACAC 274
Db 21 ArgAlaValSerLeuValThrAspSerThr 30
RESULT 14
US-09-798-116-25
; Sequence 25, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekerl, Paul
```



```

; APPLICANT: Vaux, David
; TITLE OF INVENTION: NO. US20020110851A1el Polypeptides, Modulatory Agents therefor
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU P05995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 29
; TYPE: PRT
; ORGANISM: synthetic
US-09-798-116-25

Alignment Scores:
Pred. No.: 0.000151 Length: 29
Score: 137.00 Matches: 29
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.60% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1 (1-1358) x US-09-798-116-25 (1-29)
QY 248 GCAGTGTCTTGTGTAACAGATAGCAGCTTACTTCTCTCAGACCAATATGCGTTG 307
DB 1 AAlaValSerLeuValThrAspSerThrPheLeuSerGlnThrThyrAlaLeu 20
QY 308 ATTGAAGCTATTACTGAATATATCTAAG 334
DB 21 lIleGluAlaIleThrGlnTyrThrlys 29

RESULT 15
US-09-764-868-701
; Sequence 701, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT#32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 701
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-701

Alignment Scores:
Pred. No.: 0.0753 Length: 401
Score: 111.00 Matches: 87
Percent Similarity: 39.50% Conservatve: 54
Best Local Similarity: 24.37% Mismatches: 112
Query Match: 4.54% Indels: 104
DB: 9 Gaps: 19

US-09-939-293-1 (1-1358) x US-09-764-868-701 (1-401)
QY 311 CAGCTATTACTGAA-----TATACTAAGCT-----GTTTATACCTTA 349
DB 97 GlnThrIleThrLysGlnTyrTyrArgAlaGlnGlnIlePheLeuValTyrAspIle 116
QY 350 ACTTCTTTTACCGACATATACAGTTTACTTGGAAATGATTCAGAGAGGAA--- 406
DB 117 SerSer-----GluArgSerTyrGlnHisIleMetLysTyrPvalSerAspValAspGlnTyr 135
QY 407 -----GATGAAGTGTGCAGCGATCATAGAGCCAGAGCTGATGATCATCAAAACAC 460
DB 136 AlaProGlnGlyValGlnIleLysIleLeuIleGlyAsnLysAlaAspGlnGlnLysArg 155
```

```

QY 461 CAAGAGTACTTGAACGTGGAAACACACTTGGATGACTGCAGTTGGTCTTTCAGAGATGGCA 520
DB 156 Gln-----ValGly---ArgGlnGlnGly 162
QY 521 GCAGAAGCTCATATATCAACATGCGCCAGATCAGGCTGTATACCGCAGGAATACATTT 580
DB 163 GlnGlnLeuAlaLysGlnTyrGlnMetAspPheTyrGlnThrSerAlaCysThrAsnLeu 182
QY 581 CAGCTG-----GTGAACCTGCAGCTGCAGAG 607
DB 183 AsnIleLysGlnSerPheThrArgLeuThrGlnLeuValLeuGlnAlaHisArgLysGln 202
QY 608 GTGCACACAGCTCTCCCGGAAACAGAACCAACCAAGCTGCAGAGACAGATACAGAGCTC 667
DB 203 LeuGlnGlyLeuArgMetArgAlaSerAsnGlnLeuAlaLeuAlaGlnLeuGlnGln--- 221
QY 668 CGTCAAAAACACACAGAGAGAGG-----GAGAGACGGCGTGAATCGGAGCAGAGAGGCC 721
DB 222 -----GlnGlnGlnLysProGlnGlnGlyProAlaAsnSerSerLysThrCys 236
QY 722 TACCTGCGTGAAGATTGAAGGCGCTG---AGCACACTGCCCTGTCTCCCACTCAG---TG 775
DB 237 Trp-CysCysValLeuTyrPglAlaProHisThrThrProLeuPheProGlnGlnAlaAr 256
QY 776 GGGAAAGCAGGGGCGAGATGCCACCTGCCAGGCTTGGCATCTGCTGTC----- 828
DB 256 gGlyGlnThrGlyGlnProGlyLeuCysProAla---AlaValAlaLeuSer-CysAspAsp 275
QY 829 --ACGAGAAAGAGGGCGGCGAGTCTGCCCTGCACATACGAGGAGACGCGCTTGTGAGCT 886
DB 275 rolleGlnTyrGlnLysProLeuLeuProLeuPro----- 286
QY 887 GTAGTGCCTCTGTGTGCTCAGGCTTGGCTGAGCT-----GATTCTTACGCCCTTG 940
DB 287 -----GlyProGlnSerGlySerAlaValIles 296
QY 941 GCACGTGCACCCCTGT-----TTACATTTACACCCA 970
DB 296 erSerSerProCysProGlnProAlaHisProGlyValGlnPhePheSerLeuPheProG 316
QY 971 CTCTGTACAGCTGCTCTTACCCATTTTCTTACCTGCACACCCAAAGCATTTTGGCTTACT 1030
DB 316 lnpProGlnAlaCysTyrAspProHisAspValProGlnAlaLeuSerHisHisProAlaP 336
QY 1031 GGGTCAGAGAGAGG---AGTCCCTTTTGTGCATGCCCTTAAGTTACGACCACTTTAACT 1087
DB 336 roThrArgGlnGlnProGly-LeuGlnSerArgProLeuSerAlaAlaPro-PheSerVa 355
QY 1088 GTTT-----TCAGTCTTATTACGTCGTCGCAAAATGATTTAGTACTTGTCCCTCT 1138
DB 355 lHisArgValSerSerLeuPheLeuSerSerProThrSerLeuSerLeuThrProPr 375
QY 1139 GTTGGAGATGCCAGTTGTGCAGAGGGGAGGAGCAACCTGTCCATTTGTACGATTTCTTGT 1198
DB 375 oleuArgCysValSerTyrGlnSerSerAsn----- 386
QY 1199 ATGATTTCTGATGTGTTCTGTGATCTGCCCCCACTGTCTGT 1241
DB 387 -----ProValProArgValSerCys 393
```

Search completed: February 20, 2003, 15:15:50
Job time : 19.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 20, 2003, 14:47:23 : Search time 21.5 Seconds

(without alignments)
5239.521 Million cell updates/sec

Title: US-09-939-293-1

Perfect score: 2446

Sequence: 1 ggcgtccgcgcgcgcacaa.....ggacttaacacagaanaaaa 1338

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+np2.model -DEV=xip
-O=/cgn2.1/USPPO/spool/US09939293/runat_20022003.111510.14829/app-query.fasta.1.1543
-DB=SWISSProt.40 -QFMT=fasta -SUFFIX=resp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=humand4.cdi -LIST=45
-DOCALLIG=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09939293_@CGN.1.1.30.@runat_20022003.111510.14829 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1201	49.1	239	1	SMAC_HUMAN
2	1016	41.5	237	1	SMAC_MOUSE
3	125.5	5.2	502	1	WASP_HUMAN
4	112	4.6	360	1	DMD_CHICK
5	111	4.6	1323	1	NME4_MOUSE
6	111	4.6	1323	1	NME4_MOUSE
7	110	4.5	589	1	SPY_DROME
8	107.5	4.4	1520	1	TOP2_CAEEL
9	106.5	4.4	3164	1	TB0U_CAEEL
10	106	4.4	520	1	WASP_MOUSE
11	104.5	4.3	1219	1	JAG1_RAT
12	103.5	4.2	4684	1	PLE1_HUMAN
13	103	4.2	592	1	LAM2_MOUSE
14	103	4.2	1433	1	RST_CHICK
15	102.5	4.2	3924	1	ANK2_HUMAN
16	102.5	4.2	5376	1	ZAN_MOUSE
17	102	4.2	2116	1	MYS2_DICDI
18	101.5	4.2	500	1	FXE2_HUMAN

19	101	4.1	502	1	K2M3_SHEEP
20	101	4.2	1509	1	GSRL_HUMAN
21	100	4.1	4377	1	ANK3_HUMAN
22	99.5	4.1	1087	1	NFR_MOUSE
23	99.5	4.1	1213	1	FNX_CHICK
24	99	4.1	616	1	REPX_HUMAN
25	99	4.1	721	1	YR82_MYCTU
26	99	4.1	1336	1	NME4_HUMAN
27	98.5	4.1	988	1	CIC1_HUMAN
28	98	4.0	548	1	NFL_PIG
29	98	4.0	4427	1	PKSL_BACSU
30	97.5	4.0	975	1	CUT1_CANFA
31	97.5	4.0	4473	1	PLE1_CRIGR
32	97	4.0	603	1	NMEL_MOUSE
33	97	4.0	831	1	NFR_RAT
34	97	4.0	4687	1	PLE1_RAT
35	96.5	4.0	189	1	AMEL_PIG
36	96.5	3.9	331	1	MAN1_MOUSE
37	96.5	3.9	651	1	VU47_HSV6U
38	96.5	3.9	1043	1	FPT1_DROME
39	96	4.0	633	1	LAI7_YEAST
40	96	4.0	1183	1	DRPL_RAT
41	95.5	3.9	460	1	MYCN_MARMO
42	95	3.9	544	1	NFL_XENLA
43	95	3.9	817	1	VRPL_YEAST
44	95	3.9	1208	1	RCQ4_HUMAN
45	94.5	3.9	503	1	WALP_HUMAN

ALIGNMENTS

RESULT 1

SMAC_HUMAN STANDARD: PRT: 239 AA.

ID Q9NR28; Q9NR28; Q9HAV6; Q96LW0;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Smac protein, mitochondrial precursor (Second mitochondria-derived

activator of caspase) (Direct IAP binding protein with low pI).

GN SMAC OR DIABLO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), PARTIAL SEQUENCE, FUNCTION, AND TISSUE

RP SPECIFICITY.

RP MEDLINE=20383536; PubMed=10929711;

RA Du C., Fang M., Li Y., Li L., Wang X.;

RT "Smac, a mitochondrial protein that promotes cytochrome c-dependent

caspase activation by eliminating IAP inhibition.";

RL Cell 102:33-42(2000).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,

RA Suzuki Y., Odayashi M., Nishi T., Shibaara T., Tanaka T.,

RA Nakamura Y., Isogai T., Sugano S.;

RT "NEO human cDNA sequencing project.";

RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.

RA Pubmed=10950947;

RA Sridivasula S.M., Datta P., Fan X.J., Fernandes-Alnemri T., Huang Z.,

RA Alnemri E.S.;

RT "Molecular determinants of the caspase-promoting activity of

Smac/DIABLO and its role in the death receptor pathway.";

RL J. Biol. Chem. 275:36152-36157(2000).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RA Tissue=Cerebellum;

RC Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,

RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,

RA Q99526 homo sapien

RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
 RA Kawakami B., Nagai K., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Muscle, and Uterus;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 56-239.
 RA MEDLINE-20426096; PubMed-10972280;
 RT Chai J., Du C., Wu J.W., Kyn S., Wang X., Shi Y.;
 RT "Structural and biochemical basis of apoptotic activation by
 RT Smac/Diablo."
 RL Nature 406:855-862(2000).
 RN [7]
 RP STRUCTURE BY NMR OF 56-64 IN COMPLEX WITH BIRC4.
 RA MEDLINE-21020961; PubMed-11140637;
 RT Liu Z., Sun C., Olejniczak E.T., Meadows R.P., Betz S.F., Cost T.,
 RA Herrmann J., Wu J.C., Fesik S.W.;
 RT "Structural basis for binding of Smac/Diablo to the XIAP BIR3
 RT domain."
 RL Nature 408:1004-1008(2000).
 CC -1- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/Apaf-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
 CC INHIBITORY ACTIVITY OF APOPTOSIS PROTEINS (IAP).
 CC -1- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
 CC BIRC7.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
 CC WHEN CELLS UNDERGO APOPTOSIS.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/Diablo-S;
 CC are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: UNIDIRECTIONALLY EXPRESSED WITH HIGHEST EXPRESSION
 CC IN TESTIS. EXPRESSION IS ALSO HIGH IN HEART, LIVER, KIDNEY,
 CC SPLEEN, PROSTATE AND OVARY. LOW IN BRAIN, LUNG, THYMUS AND
 CC PERIPHERAL BLOOD LEUKOCYTES.
 CC -1- DOMAIN: The mature N-terminus mediates interaction with
 CC BIRC4/XIAP.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF262240; AAR87716.1; -
 DR EMBL: AK024768; BAB14994.1; -
 DR EMBL: AF298770; AAG22077.1; -
 DR EMBL: AK057778; BAB71568.1; -
 DR EMBL: BC004417; AAH04417.1; -
 DR PDB: 1FEW; 13-SEP-00.
 DR PDB: 1G3F; 10-JAN-01.
 DR MIM: 605219; -
 KM Transl. peptide: Mitochondrion; Apoptosis; Alternative splicing;
 KM 3D-structure.
 FT TRANSIT 1 55 MITOCHONDRION.
 FT CHAIN 56 239 SMAC PROTEIN.
 FT SITE 56 60 IAP-BINDING MOTIF (BY SIMILARITY).
 FT VARSPLIC 1 60 MAALKMSLSRSTSTFFPRROCLCPVYANFKKRCFSELIRP
 FT WHKLTGSGVGLCAVPIA -> MKSDPEYF (IN
 FT ISOFORM 2).
 FT K -> E (IN REF. 4).
 FT K -> R (IN REF. 2).
 FT MISSING (IN REF. 4).
 FT E -> K (IN REF. 4).
 FT OX
 SQ SEQUENCE 239 AA; 27131 MW; 70C2AE0DC654D031 CMC64;
 Alignment Scores: 1.51e-80 Length: 239
 Pred. No.: 1201.00 Matches: 239
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 49.10% Indels: 0
 DB: 1 Gaps: 0
 US-09-939-293-1 (1-1358) x SMAC_HUMAN (1-239)
 QY 20 ATGGCGCTCTGAGAGAGTGGCTGCGCAGCCGTAATCTTCATCTTCAGTAGACAG 79
 DB 1 MetAlaIaLeuLysSerThrLeuSerArgSerValThrSerPhePheArgTyrArgGln 20
 QY 80 TGTGTGTGTGTCTCTGTTGGCTTAATCTTAAAGACGGCTTTCTCAGAAATTGATAGA 139
 DB 21 CysLeuCyValProValAlaAspPheLysArgCysPheSerGlnLeuLeuArg 40
 QY 140 CCATGGCAAAATCTGATGAGATGGCTTTGAGAAACCGTGGTGGCTTCATTTACA 199
 DB 41 ProThrHisLysThrValThrIleGlyPheGlyValThrLeuCysAlaValProIleAla 60
 QY 200 CAGAAATCAGAGCCCTATTCCTTAGTAGTAGAGCATTTGATGAGAGACAGTCTTTG 259
 DB 61 GlnLysSerGlnProHisSerLeuSerSerGlnAlaLeuMetArgAlaValSerLeu 80
 QY 260 GTACAGATACACCTCTACCTTTCTCTCTCAGACACATATGCGTGTATGAACTATT 319
 DB 81 ValThrAspSerThrSerThrPheLeuSerGlnThrTyrAlaLeuIleGlnAlaIle 100
 QY 320 ACTGAATCTACTAAGGCGTTTATACCTTACTCTCTTCCGCAATATACAGTTTA 379
 DB 101 ThrGlnTyrThrLysAlaValTyrThrLeuThrSerLeuTyrArgGlnTyrThrSerLeu 120
 QY 380 CTGGGAAATATGATTCAGAGAGAGAGATGAGTGGCAGGTGATCATAGAGCCAGA 439
 DB 121 LeuGlyLysMetAsnSerGlnGlnLysPglValTyrGlnValIleIleGlyAlaArg 140
 QY 440 GCTGATGATCTTCAAAACACCAAGACACTTGAAGCTGGAACCACTTGGATGATGCA 499
 DB 141 AlaGlnMetThrSerLysHisGlnGlnTyrLeuLysLeuGlnThrThrMetThrAla 160
 QY 500 GTTGCTCTTTCAGAGATGAGCAGAGCAAGCTCATATCAATCTGGCGCAGATCAGCCCTCT 559
 DB 161 ValGlyLeuSerGlnMetAlaAlaGlnAlaAlaIleTyrGlnThrGlyAlaAspGlnHisSer 180
 QY 560 ATAAACCGCAGGAATACATTCAGCTGCTGTAACCTGAGTGGAGAGAGTGCACCAGCTC 619
 DB 181 IleThrAlaArgAsnHisIleGlnLeuValLysLeuGlnValGlnGlnValHisGlnLeu 200
 QY 620 TCCCGGAAGCAGAAACCAACAGCTGCGACAGACACAGATAGAAAGCTCCGTCAGAAAACA 679
 DB 201 SerArgLysAlaGlnThrLysLeuAlaGlnAlaGlnIleGlnGlnLeuArgGlnLysThr 220
 QY 680 CAGGAGAGAGGAGGAGAGCGGCTGAGTCGAGAGCAGAGAGGCTACCTCGTAGAGAT 736
 DB 221 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 239
 RESULT 2
 ID SMAC_MOUSE STANDARD: PRT: 237 AA.
 AC Q9TIO3; Q9ICD1; Q9ICD3.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Smac protein, mitochondrial precursor (Second mitochondria-derived
 DE activator of caspase) (Direct IAP binding protein with low PI).
 GN SMAC OR DIABLO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RP STRAIN=BAIB/c; TISSUE=Kidney;

RX MEDLINE-20383537; PubMed=10929712;
 RA Veragen A.M., Ekert P.G., Pakusch M., Silke J., Connolly L.M.,
 RA Reid G.E., Moritz R.L., Simpson R.J., Vaux D.L.;
 RT Identification of DIABLO, a mammalian protein that promotes apoptosis
 RT by binding to and antagonizing IAP proteins.";
 RL Cell 102:43-53(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE-21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa I., Saito K.,
 RA Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakai K., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyono K., Wang K.H., Welter C., Whitaker C.,
 RA Widing L., Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H.,
 RA Kohsaki S.;
 RT Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/APAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
 CC INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
 CC -1- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
 CC BIRC7 (By similarity).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
 CC WHEN CELLS UNDERGO APOPTOSIS.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION FOUND IN HEART, LIVER,
 CC KIDNEY AND TESTIS.
 CC -1- DOMAIN: The mature N-terminus mediates interaction with
 CC BIRC4/XIAP (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF203914; AAF82190.1; -;
 DR EMBL: AK012760; BAB28450.1; -;
 DR EMBL: AK002887; BAB22433.1; ALT_FRAME.
 DR HSSP: Q9NR28; 1FEW.
 DR MGD: MGI:1913843; 0610041G12R1k.
 KW Transit peptide; Mitochondrion; Apoptosis.
 FT TRANSIT 1 53 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 1 237 SMAC PROTEIN.
 FT SITE 54 58 IAP-BINDING MOTIF (BY SIMILARITY).
 FT CONFLICT 64 64 H -> Q (IN REF. 2).
 SQ SEQUENCE 237 AA; 26829 MW; E53E6F04F1C390A1 CRC64;

 Alignment Scores:
 Pred. No.: 5,35e-67 Length: 237
 Score: 1016.00 Matches: 204
 Percent Similarity: 91.21% Conservative: 14
 Best Local Similarity: 85.36% Mismatches: 19
 Query Match: 41.54% Indels: 2
 DB: 1 Gaps: 1
 US-09-939-293-1 (1-1358) x SMAC_MOUSE (1-237)
 QY 20 ATGGCGGCTCTGTAAGAGTGGCTGTGGCGACGCTAACTTCATTCCTTCAGGTACAGACAG 79

DB 1 MetAlaAlaLeuArgSerTrpValThrArgSerValCysSerLeuHeuArgTrpArgGln 20
 QY 80 TGTGTGTGTGTCTCTGTTGGCTTAACCTTAAAGACGGGTGTTCTCAGAAATGATGAAGA 139
 DB 21 ArgPhe-----ProValLeuAlaAsnSerLysArgCysPheSerGluLeuIleLys 38
 QY 140 CCATGCGACAAAACACTGTGACGATGGCTTTGGAGTACCCCTGTGCGGTCTCTATTGCA 139
 DB 39 ProTrpHisLysThrValLeuThrGlyPheGlyMetThrLeuLysAlaValProIleAla 58
 QY 200 CAGAATTCAGAGCCCTATTCCCTTACTAGTGAAGCATGTTGATGGAGACAGCTGTTTG 259
 DB 59 GlnLysSerGluProHisSerLeuSerAsnGluAlaLeuMetLysArgAlaValSerLeu 78
 QY 260 GTAACGATAGACACCTCTACCTTCTCTCTCAGACACCATATGCTGTGATGAAGTATT 319
 DB 79 ValThrAspSerThrSerThrPheLeuSerGlnThrThrTyralaLeuIleGlnAlaIle 98
 QY 320 ACTGAATATATCTAAGCTGTTTATACCTTACTTCTTACCTTACCGACAATATACAACTTA 379
 DB 99 ThrGluTrpThrLysAlaValLysThrLeuValSerLeuTrpArgGlnTrpThrSerLeu 118
 QY 380 CTGGGAAATGATGATCTCAGAGAGAGAGATGAAGTGAAGTGGCAGGCTGATCAGAGCCAGA 439
 DB 119 LengLysMetLysSerGlnGlnLysAlaValTrpGlnValIleIleGlnAlaVal 138
 QY 440 GCTGAGATGATCTCAAAACACCAAGACTCTTGAACCTGGAACCACTGGATGATGCTGA 499
 DB 139 ValGluMetThrSerLysGlnGlnLysLysLysLysLysLysLysLysLysLysLys 158
 QY 500 GTTGGCTTTTACAGATGCGCAGAGAGAGTGCATATCAAACTGGCCAGATCAGGCTCT 559
 DB 159 ValGlyLeuSerGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 178
 QY 560 ATACCGCCAGCAATCATATCTAGCTGTTGAACCTCAGTGGAGAGAGTGGCAGGCTCT 619
 DB 179 IleThrAlaArgAsnHisIleGlnLeuValLysSerGlnValGlnLysLysLysLys 198
 QY 620 TCCCGGAAAGCAGAAACCAAGCTGSCAGAGACAGATGAGACAGCTCCGTCAGAAACA 679
 DB 199 SerGlnLysAlaGluThrLysLysLysLysLysLysLysLysLysLysLysLysLys 218
 QY 680 CAGGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 736
 DB 219 GlnGluValSerAspGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 237
 RESULT 3
 WASP_HUMAN STANDARD; PRT; 502 AA.
 AC P42768; Q9UNJ9; Q9BU11;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Wiskott-Aldrich syndrome protein (WASP).
 GN WAS OR IMD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-F-cell;
 RX MEDLINE-94549367; PubMed=8069912;
 RX Derry J.M.J., Ochs H.D., Francke U.;
 RT Isolation of a novel gene mutated in Wiskott-Aldrich syndrome.";
 RL Cell 78:635-644(1994).
 RN [2]
 RP ERRATUM.
 RX MEDLINE-95094263; PubMed=8001129;
 RX Derry J.M.J., Ochs H.D., Francke U.;
 RL Cell 79:923-923(1994).
 RN [3]

RP SEQUENCE FROM N.A.
 RX MEDLINE-95273432; PubMed-7753869;
 RA Kwan S.-P., Hagemann T.L., Radtke B.E., Blaese R.M., Rosen F.S.;
 RT "Identification of mutations in the Wiskott-Aldrich syndrome gene and
 RT characterization of a polymorphic dinucleotide repeat at DXS6940,
 RT adjacent to the disease gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:4706-4710(1995).
 [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99167346; PubMed-10066431;
 RA Hagemann T.L., Kwan S.-P.;
 RT "The identification and characterization of two promoters and the
 RT complete genomic sequence for the Wiskott-Aldrich syndrome gene.";
 RL Biochem. Biophys. Res. Commun. 256:104-109(1999).
 [15]
 RP SEQUENCE FROM N.A.
 RA Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
 RA Meindl A., Rosenthal A.;
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 [16]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 [17]
 RP VARIANTS WAS K-31; M-75; P-82; C-86; H-86; C-97; K-133; E-476 AND
 RP H-30 DEL.
 RX MEDLINE-96133285; PubMed-8528198;
 RA Kolluri R., Shehabeldin A., Peacocke M., Lamhonwah A.-M.,
 RA Telchett-Kuliszewska K., Weissman S.M., Simunovich K.A.;
 RT "Identification of WASP mutations in patients with Wiskott-Aldrich
 RT syndrome and isolated thrombocytopenia reveals allelic heterogeneity
 RT at the WAS locus.";
 RL Hum. Mol. Genet. 4:1119-1126(1995).
 [18]
 RP VARIANTS WAS F-27; I-48; M-75; L-86; H-86; K-131; C-187 AND K-477.
 RX MEDLINE-96133286; PubMed-8528199;
 RA Derry J.M.J., Kerns J.A., Weinberg K.I., Ochs H.D., Volpini V.,
 RA Estivill X., Walker A.P., Francke U.;
 RT "WASP gene mutations in Wiskott-Aldrich syndrome and X-linked
 RT thrombocytopenia.";
 RL Hum. Mol. Genet. 4:1127-1135(1995).
 [19]
 RP VARIANTS WAS VAL-56 AND GLU-236.
 RX MEDLINE-95315993; PubMed-7795648;
 RA Villa A., Notarangelo L., Macchi P., Mantuano E., Cavaign G.,
 RA Brugnoni D., Strina D., Patrosso M.C., Ramenghi U., Sacco M.G.,
 RA Ugazio A., Vezzoni P.;
 RT "X-linked thrombocytopenia and Wiskott-Aldrich syndrome are allelic
 RT diseases with mutations in the WASP gene.";
 RL Nat. Genet. 9:414-417(1995).
 [10]
 RP VARIANTS WAS TRP-43; MET-45; MET-75 AND CYS-86.
 RX MEDLINE-97272113; PubMed-9126958;
 RA Remold-O'Donnell E., Cooley J., Shcherbina A., Hagemann T.L.,
 RA Kwan S.-P., Kenney D.M., Rosen F.S.;
 RT "Variable expression of WASP in B cell lines of Wiskott-Aldrich
 RT syndrome patients.";
 RL J. Immunol. 158:4021-4025(1997).
 [11]
 RP VARIANTS WAS LYS-31 AND MET-45.
 RX MEDLINE-97253413; PubMed-9098856;
 RA Ariga T., Yamada M., Sakiyama Y.;
 RT "Mutation analysis of five Japanese families with Wiskott-Aldrich
 RT syndrome and determination of the family members' carrier status
 RT using three different methods.";
 RL Pediatr. Res. 41:535-540(1997).
 [12]
 RP VARIANTS WAS MET-75; LEU-84; ASP-89 AND LYS-133.
 RX MEDLINE-98350091; PubMed-9683546;
 RA MacCarthy-Morrogh L., Gaspar H.B., Wang Y.-C., Katz F., Thompson L.,
 RA Layton M., Jones A.M., Kinnon C.;
 RT "Absence of expression of the Wiskott-Aldrich syndrome protein in

RT peripheral blood cells of Wiskott-Aldrich syndrome patients.";
 RL Clin. Immunol. Immunopathol. 88:22-27(1998).
 [13]
 RP VARIANTS WAS VAL-56.
 RX MEDLINE-96378988; PubMed-9713366;
 RA Facchetti F., Blanzuoli L., Verini W., Notarangelo L.D., Gilliani S.,
 RA Fiorini M., Fasth A., Stewart D.M., Nelson D.L.;
 RT "Defective actin polymerization in EBV-transformed B-cell lines from
 RT patients with the Wiskott-Aldrich syndrome.";
 RL J. Pathol. 185:99-107(1998).
 [14]
 RP VARIANTS WAS LYS-133.
 RX MEDLINE-98092074; PubMed-9445409;
 RA Parolini O., Resemann G., Haas O.A., Pawlowsky J., Gadner H.,
 RA Knapp W., Holter W.;
 RT "X-linked Wiskott-Aldrich syndrome in a girl.";
 RL New Engl. J. Med. 338:291-295(1998).
 [15]
 RP VARIANTS WAS ARG-73; MET-75; CYS-83; CYS-86 AND LYS-133.
 RX MEDLINE-99374925; PubMed-10447259;
 RA Lemahieu V., Gastier J.M., Francke U.;
 RT "Novel mutations in the Wiskott-Aldrich syndrome protein gene and
 RT their effects on transcriptional, translational, and clinical
 RT phenotypes.";
 RL Hum. Mutat. 14:54-66(1999).
 [16]
 RP VARIANTS WAS HIS-52 AND TRP-70.
 RX MEDLINE-21652748; PubMed-11793485;
 RA El-Hakeh J., Rosenzweig S., Oleastro M., Bassack N., Berozdink L.,
 RA Molina F., Rivas E.M., Zelazko M., Danielian S.;
 RT "Wiskott-Aldrich syndrome in Argentina: 17 unique, including nine
 RT novel, mutations.";
 RL Hum. Mutat. 19:186-187(2002).
 [17]
 RP FUNCTION: POSSIBLE REGULATOR OF LYMPHOCYTE AND PLATELET FUNCTION.
 MAY BE INVOLVED IN SIGNALING PATHWAYS WITH CYTOSKELETAL FUNCTION.
 CC -1 SUBUNIT: BINDS TO CDC42, RAC, NCK, FYN, SRC KINASE FGR, BTK, ABL,
 CC WIP, AND TO THE P85 SUBUNIT OF PLC-GAMMA.
 CC -1 TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE THYMUS. ALSO
 CC FOUND TO A MUCH LESSER EXTENT, IN THE SPLEEN.
 CC -1 DOMAIN: THE WH1 (WASP HOMOLOG 1) DOMAIN MAY BIND A PRO-RICH
 CC LIGAND.
 CC -1 DISEASE: DEFECTS IN WAS ARE THE CAUSE OF WISKOTT-ALDRICH SYNDROME
 CC (WAS), AN X-LINKED RECESSIVE IMMUNODEFICIENCY CHARACTERIZED BY
 CC ECZEMA, THROMBOCYTOPENIA, RECURRENT INFECTIONS, AND BLOODY
 CC DIARRHEA. DEATH USUALLY OCCURS BEFORE AGE 10.
 CC -1 DISEASE: DEFECTS IN WAS ARE THE CAUSE OF ISOLATED X-LINKED
 CC THROMBOCYTOPENIA (XLT). XLT IS CLINICALLY MILD WITH SMALL
 CC PLATELETS AND SUBCLINICAL LEUKOCYTE ABNORMALITIES.
 CC -1 SIMILARITY: CONTAINS 1 CRIB DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 WH1 DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U12707; AAA62663.1; ALT_SEQ.
 CC EMBL: U18935; AAA60381.1; -.
 CC EMBL: U19927; AAC50140.1; -.
 CC EMBL: AF115549; AAD26691.1; -.
 CC EMBL: AF196970; AAR05804.1; -.
 CC EMBL: BC002961; AAH02961.1; ALT_INIT.
 CC EMBL: BC012738; AAH12738.1; -.
 CC Genew: HGNC:12731; WAS.
 CC MIM: 300392; -.
 CC MIM: 301000; -.
 CC InterPro: IPR000095; PAKbox/Rhobinding.
 CC InterPro: IPR000697; RanBP1_WASP.
 CC InterPro: IPR001960; WH1.
 CC InterPro: IPR003124; WH2.

Alignment Scores:	
Pred. NO.:	0.0966
Score:	125.50
Percent Similarity:	35.008
Best Local Similarity:	26.438
Query Match:	5.208
DB:	1
Length:	502
Matches:	74
Conservative:	24
Mismatches:	102
Indels:	80
Gaps:	13

```

      2.200  INCLERS:  00
DB: 1      GAPS: 13

```

OY	1291	CCTTCCTGTAGTGAATGATTGGTT---TTACAGCCCTCGGCCTCAGCTGCCCTCCACAGAC	1235
Dd	205	ProAspIlethrSerIserAlaGlytLAyagGlyLeuProIalProGIYProSerProIalasp	224
OY	1234	-----AGTGGGGCAGCATCAGAGAACAACATCAGAAATACATACAAGAAATCGTAC	1184
Dd	225	LysLysArgSerGlyLysLysLysIleSerLysAlaAspIIlgLYAlaProSerGlyphe	244
OY	1183	AAATGACACAGGTTCCTCCCTCCCCTGCCAC---	AACTGGCATCCCAACAGA 1138
Dd	245	Lys-----HisValSerHisValGIYTPraSprProGlnasn	256
OY	1135	GCGAACACAGTACTAATATCATTTTTTGACGACAGTAATATTAAGTCAAACAGCTTAAACAGT	1076
Dd	257	GLY-----PheAspValAlaAsnIleuAspProAspIleuArgSer	269
OY	1075	TGCTGAACCTTAAGGGCAGCATGACAAAAAGACATCCTCTCTGTGACCAGTAGGCAAAATGC	1016
Dd	270	LeuPheSerArgIalGlyIleSerGlu-----AlaIleu	281
OY	1015	TTTGGCTGTAGCGTAAAAAAAATGGGTAAAGACAGCTGTACAGAGTGGGT-----	965
Dd	282	ThrAspIalIuThrSerLysLeuIleTYrAspPheIleGluAspGlnGlyLeuGlu	301
OY	964	-----GAAATGTTAAACAGGGCAGTGGCCCA-----	AGG 935
Dd	302	AlaValArgIngluMetArgAlaGlyIngluIProLeuProProProProProProSerArg	321
OY	934	GCTNAGACACAGGTCCAGCGCAAGCCTGAGACCACAGGAGGAGCTACACAGTCAACAAAG	875
Dd	322	GIYGIYAsnIleuProArg---ProproIleValGIYGVASnuys-----	336

QY	874	CGTCTGCGCTGATGTCGCGAGGCGAGCACT	-----GCCGCC	839
Db	337	-----GlyArgSerGlyProLeuProProValProLeuGlyIleAlaPro	111:::111111	351
QY	838	TCTTCTCGGTGCACAGACAGATCATTCCCAACCTGGCGAGGCGGATGCTGCCCTGCTT	11111111111111111111	779
Db	352	ProProProthProArgGlyProProProProGlyArgGlyGlyProProProPro	11111111111111111111	371
QY	778	CCCCCATGATGGAGGACAGACAGAGTGTGCTGACAGCCCTCAATCCATCCAGCAGGTAGC	11111111111111111111	719
Db	372	ProProAlaIthGlyArgSerGlyProLeuProProProPro--ProIlyAlaGlyGlyP	11111111111111111111	391
QY	718	CTCTGCTCGCAGCTACAGCCGCTCCTCCCTTCCTCTGTGTGTTCTGACGAGACTCTTC	11111111111111111111	659
Db	391	roProMetProProProProProProProProProProSerSerGly	11111111111111111111	407
QY	658	TATCTGTCTTCTGCGACACTGTGTTCTGCTTCGGAGAGCTGTGACACTCTTCAC	11111111111111111111	599
Db	408	-----AsnGlyProAlaProProProLeuProP	11111111111111111111	417
QY	598	CTCGAGTTTCACGAGCTGAA--GTGATTCCTGGCGGTATTAAGAGCGCTATC	11111111111111111111	548
Db	417	roAlaLeuValProAla--GlyGlyLeuAlaProGlyGlyGlyAlaGlyAlaLeu	11111111111111111111	434
RESULT 4				
DM	DM	CHICK	STANDARD;	PRT; 3660 AA.
AC	01-OCT-1989	(Rel. 12, Created)		
DT	01-OCT-1989	(Rel. 12, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Dystrophin.			
GN	DMD.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=69098331; PubMed=3062582;			
RA	Lemaire C., Heilig R., Mandel J.L.;			
RT	"The chicken dystrophin cDNA: striking conservation of the C-terminal			
RL	Nucleotide sequence of chicken dystrophin cDNA."			
RL	Nucleic Acids Res. 16:11815-11815(1988).			
CC	[2]			
CC	SEQUENCE FROM N.A.			
CC	TISSUE=Muscle;			
CC	MEDLINE=69210800; PubMed=3072195;			
CC	Lemaire C., Heilig R., Mandel J.L.;			
CC	"The chicken dystrophin cDNA: striking conservation of the C-terminal			
CC	coding and 3' untranslated regions between man and chicken."			
CC	EMBO J. 7:4157-4162(1988).			
CC	-1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE			
CC	PLASMA MEMBRANE.			
CC	-1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY			
CC	ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,			
CC	ABP-120, ABP-180, OR BETA-FORIN).			
CC	-1- SIMILARITY: CONTAINS 2 CALPODIN-HOMOLOGY (CH) DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.			
CC	-1- SIMILARITY: CONTAINS 1 WW DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 ZZ-TYPE ZINC FINGER.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X13369; CA31746.1; -			
DR	PIR; S02041; S02041.			
DR	HSSP; P46939; IIBD.			

```

DR InterPro: IPR001589; Actbind_actuin.
DR InterPro: IPR001715; Calponin-like.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR002349; MW.
DR InterPro: IPR001202; MW_Rsp5_WWP.
DR InterPro: IPR000433; ZnF_ZZ.
DR Pfam: PF00307; CH_2.
DR Pfam: PF00397; WW_1.
DR Pfam: PF00435; spectrin_22.
DR Pfam: PF00569; ZF_1.
DR PRINTS: PR00403; WMDOMAIN.
DR SMART: SM00033; CH_2.
DR SMART: SM00150; SPEC; 21.
DR SMART: SM00456; WW; 1.
DR SMART: SM00291; ZnF_ZZ; 1.
DR PROSITE: PS00019; ACTININ_1; 1.
DR PROSITE: PS00020; ACTININ_2; 1.
DR PROSITE: PS01159; WW_DOMAIN_1; 1.
DR PROSITE: PS00020; WW_DOMAIN_2; 1.
DR PROSITE: PS01357; ZF_ZZ_1; 1.
DR PROSITE: PS0135; ZF_ZZ_2; 1.
KM Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
Repeat; Zinc-finger.
FT DOMAIN 1 244 ACTIN-BINDING.
FT DOMAIN 19 123 CH 1.
FT DOMAIN 138 241 CH 2.
FT REPEAT 341 449 SPECTRIN 1.
FT REPEAT 450 558 SPECTRIN 2.
FT REPEAT 561 669 SPECTRIN 3.
FT REPEAT 721 830 SPECTRIN 4.
FT REPEAT 832 936 SPECTRIN 5.
FT REPEAT 945 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.
FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2210 2211 SPECTRIN 15.
FT REPEAT 2214 2321 SPECTRIN 16.
FT REPEAT 2472 2574 SPECTRIN 17.
FT REPEAT 2577 2683 SPECTRIN 18.
FT REPEAT 2686 2799 SPECTRIN 19.
FT REPEAT 2802 2904 SPECTRIN 20.
FT REPEAT 2928 3037 SPECTRIN 21.
FT REPEAT 3052 3085 SPECTRIN 22.
FT DOMAIN 2931 3037 WW.
FT ZN_FING 3304 3351 ZZ-TYPE.
FT VARIANT 1171 1171 MISSING.
FT VARIANT 1869 1869 Q -> H.
FT VARIANT 1885 1885 K -> R.
SQ SEQUENCE 3660 AA; 422874 MW; 85493DAF6D5B6D4A CRC64;

```

```

Alignment Scores:
Pred. No.: 1.22 Length: 3660
Score: 112.00 Matches: 56
Percent Similarity: 40.54% Conservative: 49
Best Local Similarity: 21.62% Mismatches: 112
Query Match: 4.58% Indels: 42
DB: 1 Gaps: 8

```

US-09-939-293-1 (1-1358) x DMD_CHICK (1-3660)

```

QY 5 TCCTGGCGCTGCACATGCGGCTCTGAAGAGTTGCTGCGGAGCGTAACCTTCATTC 64
DB 1163 SerLeuArgLysAspLeuSerGluMetHisGluTrpIleThrGlnAlaGluGluGluTrp 1182
QY 65 TTCAGGTACAGACAGCTGTTTGTGTGTCTCTGTTGGCGTAACCTTTAAAGAGCGGCTTTTC 124
DB 1183 LeuGlu-----ArgAspPheGluTrpLysThrPro 1192

```

```

QY 125 TCAGATTGATTAAGACCATGCGACAAAACCTGTGAGATTGGCTTTGGATTAACCTGTGT 184
DB 1193 GluGluLeuGlnLys----- 1197
QY 185 GCGGTTCTTATTGCACAGAAATCAGAGCCCTCATTCCTTAGTAGGAACATTGATGAGG 244
DB 1198 AlaValGluGluLeuLysArgLysAlaLysGluAspAlaMetGlnLysGluVal----- 1214
QY 245 AGAGCAGTCTTTGTGTACAGATAGACACCTGACTTCTCTCTCAGACCCATATGCG 304
DB 1215 ---LysValLysLeuIleThrAspSerValAsnAspPheIleAlaLysAlaProProAla 1233
QY 305 TTGATTGAAGCTATTACTGATATATCTAAGCGCTGTTTATACCTTAACCTTCTTACCGA 364
DB 1234 AlaAsnGluAlaLeuLysLysGluLeuAspValLeu-----IleThrSerTyrGlnArg 1251
QY 365 CAATATACAAAGTTTACTTGGGAAATGAATTCAGAGGAGGAGATGAAGTGTGCGAGGTG 424
DB 1252 LeuCysSerArgLeuAsnGlyLysCysLysThrLeuGlu-----GluValTrp----- 1267
QY 425 ATCATAGAGCCAGAGCTGAGATGACTTCAAAACACCAAGAGTACTTGAAGCTGGAAC 484
DB 1268 -----AlaCysTrpHisGluLeuLeuSer-----TyrLeuAspAlaGluAsn 1281
QY 485 ACTTGATGACTGTCAGATTGGTCTTTCAGAGATGCGCAGAGAGCTGATATCAACTGGC 544
DB 1282 LysTrpLeuAsnGluValGluLeuLysLeuLysAlaIleTrpGlu---AsnIleGlnLysGly 1300
QY 545 GCAGATCAGGCGCTCTATACCGCCAGATCATCATTCAGCTGCTGTAACCTGCGAGGTGGA 604
DB 1301 AlaGluGluIleSerGluSerLeuAspSerLeuGluArgLeuMetArgHisProGluAsp 1320
QY 605 GAGGTGCACCCACTCTCCCGGAAGACGAACCAAGCTGGCGGAGACAGATTAAGAG 664
DB 1321 AsnArgAsnGlnIleArgGluLeuAlaGlnTrpIleThrAspGlyLysIleLeuAspGlu 1340
QY 665 CTCCTCCTCAAAAACACAGAGGAGGAGGAGCGGCTGATCGCAGAGAGGCTCTAC 724
DB 1341 LeuIleAsnGluLysLeuGlnLysPheAsnTrpArgTrpGluGlnLeuGlnIleGlnAla 1360
QY 725 CTGCTGAGAGATTGAGGCGCTGAGCAGACACTGCCCTGTCTCCCGCATGAGGGGAAA 781
DB 1361 ValArgArgGlnLysSerLeuGlnLysIleGlnSerAlaGlnIleThrAspLys 1379

```

RESULT 5

```

ID NME4_MOUSE STANDARD; PRT; 1323 AA.
AC Q03391;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate [NMDA] receptor subunit epsilon 4 precursor (N-methyl
DE D-aspartate receptor subtype 2D) (NR2D) (NMDAR2D).
GN GRIN2D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93050214; PubMed=1385220;
RA Ikeda K., Nagasawa M., Mori H., Araki K., Sakimura K., Watanabe M.,
RA Inoue Y., Mishina M.,
RA "Cloning and expression of the epsilon 4 subunit of the NMDA receptor
RT channel.";
RL FEBS Lett. 313:34-38(1992).
RP [2]
RP REVISIONS.
RA Ikeda K., Nagasawa M., Mori H., Araki K., Sakimura K., Watanabe M.,
RA Inoue Y., Mishina M.;
RA Submitted (May-1998) to the EMBL/GenBank/DBJ databases.

```


Db	1044	AlahbheluaspGlnSerProProAlaProSerAArgTrpProArgSerAspProGlnSer	1063
Qy	765	-----GGAGACAGGCGAGTGTCGACGGCCCTCAATCCTCAGCAGGTAGGCC	718
Db	1064	GlnProLeuIeuGlyGlyGlyAla-----GlyGlyProSerAlaGly-AlaProThrAl	1081
Qy	717	TTCTGCTCCGACATCAGCCCGCTCTCC-----CTTCTCTCTGTGT	676
Db	1081	apProProAlaArgThrAlaProProProCysAlaTrpLeuAspLeuGlnProSerP	1101
Qy	675	TTTCGAC-----GGAGCTCTTATCTGNCCTTCGCGACG	640
Db	1101	oSerAspSerGlnAspSerGlnSerIeuGlyAlaIleAspLeuGlyIleuGlnProTr	1121
Qy	639	TTGGTTTCTG-----CTTCCGGAGAGCGTGTGCACCP-----	606
Db	1121	pTrpPheAlaAspPheProTrpTrpTrpAlaGlnAtrgLeuGlyProProGlyAArgTy	1141
Qy	605	-----CTTCCACCTTCG	595
Db	1141	rTrpSerValAspPheLeuGlyIleTrpArgAlaGlySerTrpAspTyLeuProProAr	1161
Qy	594	AGTTTACACAGCT-----GAGTGTATTCCTCGCGGTATTAAGCCTGATCT	547
Db	1161	gGlyGlyProAlaTrpHisCysArgHisCysAlaSerIeuGlnIleuLeuProProAr	1181
Qy	546	GCGCACGTTTGATATWGCAGCTCTGCTGCCATCTCGAAGACCACTGCAGTCATCCA	487
Db	1181	gHis-----LeuSerCysSerHisAspGlyLeuAspGlyGlyTrp	1194
Qy	486	GTGCTTCCA 477	
Db	1194	pTrpAlaPro 1197	
RESULT 6			
ID	NME4_RAT	STANDARD;	PRT; 1323 AA.
AC	Q62645; Q63381; Q63382; Q63729; Q63730;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Glutamate [NMDA] receptor subunit epsilon 4 precursor (N-methyl		
DE	D-aspartate receptor subtype 2D) (NR2D) (MMDARR2D).		
GN	GRIN2D.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 1265-1356 FROM N.A.		
RP	(ISOFORM 2).		
RC	STRAIN=Sprague-Dawley; TISSUE=Forebrain;		
RX	MEDLINE=93155102; PubMed=8428955;		
RA	Isihai T., Moriyoshi K., Sugihara H., Sakurada K., Kadotani H.,		
RA	Yokoi M., Nakazawa C., Shigemoto R., Mizuno N., Masu M., Nakanishi S.;		
RT	"Molecular characterization of the family of the N-methyl-D-aspartate		
RT	receptor subunits.";		
RL	J. Biol. Chem. 268:2836-2843(1993).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Brain;		
RX	MEDLINE=94206533; PubMed=7512349;		
RA	Monyer H., Burnashev N., Laurie D.J., Sakmann B., Seeburg P.H.;		
RT	"Developmental and regional expression in the rat brain and functional		
RT	properties of four NMDA receptors.";		
RL	Neuron 12:529-540(1994).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;		
RA	Boulter J., Pecht G.,		
RT	Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.		
CC	-I- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS		
CC	WITH HIGH CALCIUM PERMEABILITY AND VOLTAGE-DEPENDENT		

RT "Sproutly encodes a novel antagonist of EGF signalling that patterns
RT apical branching of the Drosophila airways.",
RL Cell 92:253-263(1998).
RN [2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RC
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zheng Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abdl J.F., Agayanni A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencs P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke S., Dangen Z., Davenport L.B., Davies P.,
RA de Paulos B., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlitta S., Fleischman W.,
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodde C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostalko D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski R.P., Smith T.,
RA Spier E.F., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Styvras R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [3]
RN
RP FUNCTION;
RX MEDLINE=99387981; PubMed=10457022;
RA Reich A., Sapir A., Shilo B.-Z.;
RT "Sproutly is a general inhibitor of receptor tyrosine kinase
RT signaling.";
RL Development 126:4139-4147(1999).
RN [4]
RN
RP FUNCTION;
RX MEDLINE=99244704; PubMed=10226010;
RA Kramer S., Okabe M., Hachem N., Krasnow M.A., Hiromi Y.;
RT "Sproutly, a common antagonist of FGF and EGF signaling pathways in
RT Drosophila.";
RL Development 126:2515-2525(1999).
RN [5]
RN
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=99189751; PubMed=10089881;
RA Casci T., Vinos J., Freeman M.;
RT "Sproutly, an intracellular inhibitor of Ras signaling.";
RL Cell 96:655-665(1999).
CC
CC -1- FUNCTION: INHIBITOR OF TRACHEAL BRANCHING THAT RESTRICTS BRANCH
CC BUDDING BY ANTAGONIZING THE BNL-FGF PATHWAY (BNL: BRANCHLESS, AN
CC FGF INDUCER OF BRANCHING). ACTS AS AN ANTAGONIST OF EGFR-MEDIATED
CC SIGNALING IN THE EYE (WHERE IT IS IMPORTANT FOR CELL
CC DETERMINATION) MIDLINE GLIA, CHOROIDOTONAL ORGANS, WING AND OVARIAN
CC POLARIZED CELLS.
CC
CC -1- SUBUNIT: INTERACTS WITH DRK AND GAPI PROTEIN OF THE RAS PATHWAY.
CC
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE INNER SURFACE OF THE

```

CC PLASMA MEMBRANE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYONIC TRACHEAL SYSTEM,
CC DEVELOPING EYE IMAGINAL DISK, EMBRYONIC CHORDONOTAL ORGAN
CC PRECURSORS, MIDLINE GLTA, WING DISK AND OVARY.
CC -1- DEVELOPMENTAL STAGE: FROM STAGE 7 OF OÖGENESIS FOUND IN THE
CC POSTERIOR FOLLICLE CELLS AND DURING STAGE 9 WHEN THE FOLLICLE
CC CELLS MIGRATE POSTERIORLY OVER THE OÖCYTE NUCLEUS, EXPRESSION IS
CC SEEN IN THE DORSAL AND LATERAL CELLS, AND IS EXCLUDED FROM THE
CC VENTRAL CELLS. ONCE THE MIGRATION OF FOLLICLE CELLS IS COMPLETE
CC EXPRESSED IN THE DORSAL-ANTERIOR CORNER OF THE EGG CHAMBER.
CC -1- INDUCTION: BY THE BNL-FGF PATHWAY IN THE TRACHEAL SYSTEM AND BY
CC THE EGF RECEPTOR PATHWAY IN THE WING IMAGINAL DISK AND THE
CC FOLLICLE CELLS OF THE OVARY.
CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
CC THE PROTEIN TO THE PLASMA MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AF039842; AAC04257.1; -
CC DR EMBL: AE003478; AAF47772.1; -
CC DR FlyBase: FBgn0014388; sty.
CC KW Developmental protein; Membrane.
CC FT DOMAIN 104 135 SER-RICH.
CC FT DOMAIN 148 162 ASN-RICH.
CC FT DOMAIN 207 271 GLN-RICH.
CC FT DOMAIN 378 501 CYS-RICH.
CC FT DOMAIN 492 551 GLY-RICH.
CC FT DOMAIN 249 262 POLY-GLN.
CC FT DOMAIN 503 514 POLY-GLY.
CC FT CONFLICT 38 38 N -> T (IN REF. 1).
CC FT CONFLICT 262 262 O -> Q (IN REF. 1).
CC FT CONFLICT 312 312 P -> L (IN REF. 1).
CC SO SEQUENCE 589 AA; 62388 MW; 7450E11987B17B20 CRC64;

Alignment Scores:
Pred. No.: 1.35 Length: 589
Score: 110.00 Matches: 70
Percent Similarity: 32.73% Conservative: 39
Best Local Similarity: 21.02% Mismatches: 112
Query Match: 4.50% Indels: 112
DB: 1 Gaps: 15

US-09-939-293-1 (1-1358) x SPY_DROME (1-589)
OY 518 GCAGCAGAAAGCGCATTAACAACGTGGCGCAGATCAGGCGCTATTAACGCCGACGAAATCAC 577
||| ::::: ||| ||| ::|||
Db 209 AlAtPrAySeGInHIsPrOAlaGlyInGInAsPaNgLyInThrThrHisHis 228
578 ATTGACGTGTTGAACCTCAGGTGGAAAGTGCACAGCTCTCCCGGAAGACAGAAACC 637
::: ||||::: ::|||
Db 229 LeuLeuLeuProGInAlaGAsnGInHisLeuHis----- 240
OY 638 AAGCTGGCAGAAAGCACATATGAAAGACTCCGTCAGAAAACACAGAGAGAGGAGAG 697
::: ::::: ::||| |||::: :::::
Db 241 -----LeuGInGInHisGInGInHisLeuGInGInGInGInGIn 254
OY 698 CGGGGTGAGTCGAGCAGAGAGCGCTACCTGCGTGAAGATTGAGGGCGTCAGACAC----- 751
::: ::: ||||::: ::|||:::
Db 255 GInGInGInGInGInGInGInHisLeuGInHisGInGInGInGInGInGInHisAlaArg 274
OY 752 -----ATGCGCCTGTCTCCCACTCACTGGGGAAGACAGGGCGAGAT----- 793
||| ||| ||||| |||
Db 275 LeuAlaThrThrGInAlaIaThrSerValGlySerAspHisThrAspGlyLeuLeuHis 294
OY 794 -----GCCACCTGCCCAAGGGGTGGCATATGATCATGCTGTGTCACCGACGAGA 835
||| ||| |||||

```

Db 295 SerHisLeuGlnAsnSerThrThrLysProProAlaSerLysGlnProAlaProProArg 314
 QY 836 AGAGCGGCGAGT----- 848
 Db 315 LeuGlyMetGlyLeuGlyLeuGlyLeuGlyLeuGlnProIleIleThrly 334
 QY 849 ---CTGCGCCGCGCAATCAGGCGAGAGCGCTTTGTGAGCTGTGAGTGCCTCTGTGGT 904
 Db 334 sglProThrProAlaThrGlnLysGlnArgMetHisAlaLeu-GlnGlnLeuLeuGlnP 354
 QY 905 CTCAGGCTGGGCTGAGACCTGCTTACGCCCTGGGAGACTGCA-----CCCTGTTTAA 958
 Db 354 roGlyGlyAlaGlyLeuGlnGlyGly---ProLeuValMetAlaGlyAspProSerLeu- 372
 QY 959 CATTTCACCCACCTCTG-----TACAGCTGCTCTTAC 991
 Db 373 --LeuAsnProIleValCysProArgCysGlyArgCysArgCysGlnGlnCysGlnSerP 392
 QY 992 CATTGTTTATACCTCACACCCAAAGCATTTTGCCTACCTGGGTCAGAGAGAGAGCTCTT 1051
 Db 392 roArgProLeuProGln-----ThrtProValCysAsnLysThrCysL 406
 QY 1052 TTGTGTCATGCC-----CTTAAGTTCAGACACTGTTTACCTGTTTACCTTATATT 1102
 Db 406 eucYsserAlaGlnSerValIleAspTyrAlaSerCysLeuCysCysAlaLysAlaLeu- 425
 QY 1103 TACGTGTCAAAATGATTATGACTGTGCTCTCTGTGGATGCGCATGTGGACAGG 1162
 Db 426 -----PheTyrHisCysAlaArgAspAsnAspLeuAspAspAsp 440
 QY 1163 GGAGGGAGACCTGCCAGT-----TGTACGATTTCTTTGAT- 1200
 Db 440 LysnGlyThrProCysValAspAsnProCysSerCysGlyProTyrLysArgThrGlnA 460
 QY 1201 -----GATTTCTGATG- 1212
 Db 460 rgtTpgLyTrpLeuGlyAlaLeuSerIlePheLeuProCysLeuTrpPheTyrTrpPROM 480
 QY 1213 -----TGTCTCTGATCTGCCCCCAGCTGCTGAGAGAGACGTCAGGCCAAGAGT 1264
 Db 480 eLarGlyCysMetLysLeuGlyGlnCysGlyArgGlyArgPheAla----- 495
 QY 1265 GAAAACCTATTACTACTAGAGAGAGGGGTGACAG 1299
 Db 496 -----GlyArgGlyCysArg 500
 RESULT 8
 TOP2_CABEL STANDARD; PRT; 1520 AA.
 ID TOP2_CABEL
 AC Q23670; Q27509;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Probable DNA topoisomerase II (EC 5.99.1.3).
 GN K12D12.1.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Swinburne J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Coles L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
 CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
 CC MAKES DOUBLE-STRAND BREAKS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining

CC of double-stranded DNA.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z70213; CAA94177.1; -.
 DR EMBL; Z49069; CAA94177.1; JOINED.
 DR EMBL; Z49069; CAA88867.1; -.
 DR EMBL; Z70213; CAA88867.1; JOINED.
 DR HSP; P06786; IBGM.
 DR WormPep; K12D12.1; CE06184.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR003957; CBFA_NFYB_topis.
 DR InterPro; IPR001241; DNA_topoisolI.
 DR InterPro; IPR002205; DNA_topoisolIV.
 DR Pfam; PF00204; DNA_gyraseb; 1.
 DR Pfam; PF00521; DNA_topoisolIV; 1.
 DR PRINTS; PR00615; CCAATSUBUNITA.
 DR PRINTS; PR00418; TP12FAMILY.
 DR ProDom; PD000616; DNA_topoisolI; 1.
 DR ProDom; PD000742; DNA_topoisolIV; 1.
 DR SMART; SM00433; TOP2c; 1.
 DR SMART; SM00434; TOP4c; 1.
 DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
 KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
 FT NP_BIND 196 201 ATP (POTENTIAL).
 FT ACT_SITE 840 840 DNA CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 1520 AA; 172333 MW; 016226697F41C360 CRC64;
 Alignment Scores:
 Pred. No.: 2.33 Length: 1520
 Score: 107.50 Matches: 70
 Percent Similarity: 37.42% Conservative: 49
 Best Local Similarity: 22.01% Mismatches: 110
 Query Match: 4.39% Indels: 89
 DB: 1 Caps: 15
 US-09-939-293-1 (1-1358) x TOP2_CABEL (1-1520)
 QY 92 CCGTGTGGCTAACTTAAAGACCGGTGTTCTCAGATGATTAAGACCATGGCCACAA 151
 Db 987 ProValIleValaIsprYrLysGln-----TyrHisThr 997
 QY 152 ACTGTGACGATGGCTTGGATGTAACCTGTGCGGTTCTATTGACAGAAATACGAG 211
 Db 998 AspThrThrValLysPheValValLysLeuSerProGlyLysLeuArgGlnGlnArg 1017
 QY 212 CCTATTCCTTAAGTACTAGT-----GAACATGATGAGAGAGAGAGAGTCTTTG 259
 Db 1018 GlyLlnAspLeuHisGlnValPheLysLeuGlnAlaValIleAsnThrThrCysMetVal 1037
 QY 260 GTAACAGATAGACCTTACCTTCTCTCTGACACCATATGCGTTGATTGAAGTATT 319
 Db 1038 LeuPheAspAlaAlaGlyCysLeu-----ArgThrTyrThrSerProGlnAlaIle 1054
 QY 320 ACT-----GAATATCTAAG 334
 Db 1055 ThrGlnLuphetyraspSerArgGlnGlnLysTyrValGlnArgLysGlnTyrLeuLeu 1074
 QY 335 GCTGTTTATACCTTAACCTCT-----CTTACGAGCAATATTAAGTCTGGGAAA 388
 Db 1075 GlyValLeuGlnAlaGlnSerLysArgLeuThrAsnGlnAlaArgPheIleLeuAlaLys 1094
 QY 389 ATGAATTCAGAG-----GAGAGATGAAGTGTGCGACAGTGATCATATA 430

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L38483; AAB06509.1; .

DR HSSP: P00740; 1EDM.

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR001774; DSL.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000742; EGF-2.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR001438; EGF_II.

DR Pfam: PF01414; DSL; 1.

DR Pfam: PF00008; EGF; 14.

DR PRINTS: PR00010; EGFBLD.

DR SMART: SM00051; DSL; 1.

DR SMART: SM00179; EGF_Ca; 10.

DR SMART: SM00001; EGF-like; 6.

DR SMART: SM00214; VWC; 1.

DR PROSITE: PS00010; ASX_HYDROXYL; 10.

DR PROSITE: PS00022; EGF_1; 16.

DR PROSITE: PS01186; EGF-2; 12.

DR PROSITE: PS01187; EGF_Ca; 8.

DR Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein; Repeat; Transmembrane; Signal.

KW Repeat; Transmembrane; Signal.

KM

FT SIGNAL 1 33 POTENTIAL.

FT CHAIN 1 1219 JAGGED 1.

FT DOMAIN 34 1067 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1068 1093 POTENTIAL.

FT DOMAIN 1094 1219 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 167 229 DSL.

FT DOMAIN 230 266 EGF-LIKE 1 (ATYPICAL).

FT DOMAIN 296 334 EGF-LIKE 2.

FT DOMAIN 336 372 EGF-LIKE 3.

FT DOMAIN 374 410 EGF-LIKE 4.

FT DOMAIN 412 448 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 450 485 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 487 523 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 525 561 EGF-LIKE 8.

FT DOMAIN 574 627 EGF-LIKE 9.

FT DOMAIN 629 665 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 667 703 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 705 741 EGF-LIKE 12.

FT DOMAIN 744 780 EGF-LIKE 13.

FT DOMAIN 782 818 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 820 856 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).

FT DISULFID 300 312 BY SIMILARITY.

FT DISULFID 306 322 BY SIMILARITY.

FT DISULFID 324 333 BY SIMILARITY.

FT DISULFID 340 351 BY SIMILARITY.

FT DISULFID 345 360 BY SIMILARITY.

FT DISULFID 362 371 BY SIMILARITY.

FT DISULFID 378 389 BY SIMILARITY.

FT DISULFID 383 398 BY SIMILARITY.

FT DISULFID 400 409 BY SIMILARITY.

FT DISULFID 416 427 BY SIMILARITY.

FT DISULFID 421 436 BY SIMILARITY.

FT DISULFID 438 447 BY SIMILARITY.

FT DISULFID 454 464 BY SIMILARITY.

FT DISULFID 458 473 BY SIMILARITY.

FT DISULFID 475 484 BY SIMILARITY.

FT DISULFID 491 502 BY SIMILARITY.

FT DISULFID 496 511 BY SIMILARITY.

FT DISULFID 513 522 BY SIMILARITY.

FT DISULFID 529 540 BY SIMILARITY.

FT DISULFID 534 549 BY SIMILARITY.

FT DISULFID 551 560 BY SIMILARITY.

FT DISULFID 578 605 BY SIMILARITY.

FT DISULFID 599 615 BY SIMILARITY.

FT DISULFID 617 626 BY SIMILARITY.

FT DISULFID 633 644 BY SIMILARITY.

FT DISULFID 638 653 BY SIMILARITY.

FT DISULFID 655 664 BY SIMILARITY.

FT DISULFID 671 682 BY SIMILARITY.

FT DISULFID 676 681 BY SIMILARITY.

FT DISULFID 693 702 BY SIMILARITY.

FT DISULFID 709 720 BY SIMILARITY.

FT DISULFID 714 729 BY SIMILARITY.

FT DISULFID 731 740 BY SIMILARITY.

FT DISULFID 748 759 BY SIMILARITY.

FT DISULFID 753 768 BY SIMILARITY.

FT DISULFID 770 779 BY SIMILARITY.

FT DISULFID 786 797 BY SIMILARITY.

FT DISULFID 791 806 BY SIMILARITY.

FT DISULFID 808 817 BY SIMILARITY.

FT DISULFID 824 835 BY SIMILARITY.

FT DISULFID 829 844 BY SIMILARITY.

FT DISULFID 846 855 BY SIMILARITY.

FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 559 559 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1064 1064 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1219 AA; 134325 MW; 65D4CFC38A0E204 CRC64;

Alignment Scores:

Pred. No.: 3.75 Length: 1219

Score: 104.50 Matches: 71

Percent Similarity: 29.67% Conservative: 37

Best Local Similarity: 19.51% Mismatches: 131

Query Match: 4.27% Indels: 125

DB: 1 Gaps: 19

US-09-939-293-1 (1-1358) x JAG1_RAT (1-1219)

QY 481 AACCACTGGATGACGACGCTGGTCT 507

DB 666 AsnAsnIleAsnAspCysSerGlnAsnProCysHisTyrGlyGlyThrCysArgSerLeu 685

QY 508 TTCAGAGTGGCAGCAGAAAGCTGCATATCAAACTGGCAGATCAAGCCCTATACCCG 567

DB 686 ValAsnAspPheTyrCysAspCysLysAsnGlyTyrPheGlyThrCysHisSerArg 705

QY 568 CAGGAATCATTCACCTGCTGAAACTGCAGGTGGAAGAGGTGCACCACTCTCCGGAA 627

DB 706 AspSer-----GlnCysAspGluAlaThrCysAsnAsnGlyGlyThrCysTyrArgGlu 723

QY 628 -----AGCAAAACCAAGCT 642

DB 724 ValAspThrPheLysCysMetCysProGlyGlyTyrPheGlyGlyThrCysAsnIleAla 743

QY 643 GGCAGAGCAGACAGATAGAAGACTCCGTCAGAAACACAGAGGAGGAGAGCGGC 702

DB 744 ArgAsnSerSerCysLeuProAsnProCysHisAsnGlyGly----- 757

QY 703 TGAATGCGAGCAGAGGCTACCTGCGTGAAGATTGAGGGCTGAGACACATGCCCTGTC 762

DB 758 -----ThrCysValVal 761

QY 763 TCCCCACTGATGGGGAAGCAGGGCAGATGCACCCCTGCCAGGGTGGCATGACTGT 822

DB 762 -----AsnGlyAspSer---PheThrCysValCysLysGluGlyTyrPheGluGlyPro 777

QY 823 CTGTGCACCCGAGAGAGCGCGGAGGTCTGCCCTGGCCA----- 861

DB 778 IleCysThrGlnAsnThrAsnAspCysSerProHisProCysTyrAsnSerGlyThrCys 797

QY 862 ATCAGCGCAGACGCCCTTTGTGACSTGTGAGTGCCTCTGCTGCAGGCTGGCGGTGA 921

DB 798 ValAspGlyAspAsnThrPyrArgCysGluCys-----AlaProGlyPheAlaGly 814

QY 922 CCT-----GGTCTTAGCCCTTGGGCACTGCACCCCTGTTTAACATTTCACCC 969

FT	CONFLICT	1334	1334	V > L (IN REF. 2).
FT	CONFLICT	1534	1534	I -> M (IN REF. 2).
FT	CONFLICT	1662	1662	A -> T (IN REF. 2).
FT	CONFLICT	1688	1690	WLC -> RLR (IN REF. 2).
FT	CONFLICT	1767	1767	O -> E (IN REF. 2).
FT	CONFLICT	1789	1789	L -> A (IN REF. 2).
FT	CONFLICT	1910	1910	K -> R (IN REF. 2).
FT	CONFLICT	2154	2154	K -> N (IN REF. 2).
FT	CONFLICT	2215	2215	S -> R (IN REF. 2).
FT	CONFLICT	2260	2260	R -> Q (IN REF. 2).
FT	CONFLICT	2244	2244	S -> E (IN REF. 2).
FT	CONFLICT	3027	3027	K -> E (IN REF. 2).
FT	CONFLICT	3310	3310	E -> A (IN REF. 2).
FT	CONFLICT	3361	3361	F -> L (IN REF. 2).
FT	CONFLICT	3408	3408	F -> L (IN REF. 2).
FT	CONFLICT	3447	3447	S -> A (IN REF. 2).
FT	CONFLICT	3531	3531	G -> A (IN REF. 2).
FT	CONFLICT	3580	3580	R -> S (IN REF. 2).
FT	CONFLICT	3589	3589	K -> Q (IN REF. 2).
FT	CONFLICT	3596	3596	E -> Q (IN REF. 2).
FT	CONFLICT	3616	3616	N -> H (IN REF. 2).
FT	CONFLICT	3686	3686	V -> A (IN REF. 2).

Alignment Scores:

Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

5.31
103.50
47.62%
25.85%
4.23%
1

Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:

4684
38
32
48
29
6

US-09-939-293-1 (1-1358) x PLEI_HUMAN (1-4684)

OY

386

AAATGTAATTCAAGAGGAGAAGATGAAGTGTGGCAGCTGATCATATGAGAGCCAGACTGAG

445

Db

2183

ArgLeuglnAlaetglutluySLahSLalpheAlaVal-----GlnIn

2197

OY

446

ATGACCTTCAAAACACCAGACAAGTAGTCTTGAACTGGAACCACTGGATGACTGCAGTT--

502

Db

2198

LySgLnGLnGLnLueNgInThrLaNGInGLnGLnSerValLeuAspArgLeuAryg

2217

OY

503

GGTCCTTCAAGATGGCAGACAGAACTGCATATCAAACITGGCGGCAAGATCAGGCCCTCTATA

562

Db

2218

GLyGLnALaGLnAlaAlaArGaRyArGLnALaGLnJL----AlaGLnGLnAlaRyAl

2235

OY

563

ACGCCCAAGAAATCACATTACGTGTGTGAACACAGCTGGAAGAGGTGCACCAGCTTCC

622

Db

2236

GLnAlaGLnArGLnAlaAlaGLnSerArGaRyGlnIvalIGlnGLnAlaGLnArGLnAryLeuYS

2255

OY

623

CAGAAACAGAAACCAAGCTG-----GCAGAAACAGATAGAACAGCTC

667

Db

2256

GLnSerValaGLnGLnGLnAlaGLnAlaArGaRyAlaGLnAlaGLnAlaAlaAlaGLnLysLeu

2275

OY

668

CGTCAGAAAAACACAGAGAGAGGGAGAGCGGGCTGAGCTCGAGCAGAGAGGCTTACTGT

727

Db

2276

ArgYsgLnALaGLnGLnGLnAlaAlaArGaRyArGLnAlaGLnAlaGLnAlaAla---Leu

2294

OY

728

CGTGAAGATTGAGGGCCTGAGCACACTGCCCTCTCTCCCCACTCACTGGGGAAACAGAGG

787

Db

2295

ArgGLn-----LysGLnAla

2299

OY

788

GCAGATGCCACCTGCCAGG

808

Db

2300

AlaAspAlaGLnMetGLnYS

2306

RESULT 13

LAM2_MOUSE

ID ID LAM2_MOUSE STANDARD; PRT; 592 AA.

AC P21619;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lamin B2.

CN	LAMB2.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCB1_taxid=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=g1106216; PubMed=2102682;
RA	Hoeeger T.H., Zatloukal K., Walzenegger I., Krohne G.;
RT	"Characterization of a second highly conserved B-type lamin present
RL	in cells previously thought to contain only a single B-type lamin,";
RM	Chromosome 99:379-390(1990).
RN	[2]
RP	ERRATUM.
RX	MEDLINE=g1339548; PubMed=2102440:
RA	Hoeeger T.H., Zatloukal K., Walzenegger I., Krohne G.;
RL	Chromosome 100:67-68(1990).
CC	-1 FUNCTION: LAMINS ARE COMPONENTS OF THE NUCLEAR LAMINA, A FIBROUS
CC	LAYER ON THE NUCLEOPLASMIC SIDE OF THE INNER NUCLEAR MEMBRANE,
CC	WHICH IS THOUGHT TO PROVIDE A FRAMEWORK FOR THE NUCLEAR ENVELOPE
CC	AND MAY ALSO INTERACT WITH CHROMATIN.
CC	-1 SUBCELLULAR LOCATION: NUCLEOPLASMIC SIDE OF THE INNER NUCLEAR
CC	MEMBRANE.
CC	-1 ALTERNATIVE PRODUCTS: 2 isoforms: B2 (shown here) and B3 (AC
CC	P46880); may be produced by alternative splicing.
CC	-1 PIM: B-TYPE LAMINS UNDERGO A SERIES OF MODIFICATIONS, SUCH AS
CC	FARNESYLATION AND PHOSPHORYLATION, INCREASED PHOSPHORYLATION OF
CC	THE LAMINS OCCURS BEFORE ENVELOPE DISINTEGRATION AND PROBABLY
CC	PLAYS A ROLE IN REGULATING LAMIN ASSOCIATIONS.
CC	-1 MISCELLANEOUS: THE STRUCTURAL INTEGRITY OF THE LAMINA IS STRICTLY
CC	CONTROLLED BY THE CELL CYCLE, AS SEEN BY THE DISINTEGRATION AND
CC	FORMATION OF THE NUCLEAR ENVELOPE IN PROPHASE AND TELOPHASE,
CC	RESPECTIVELY.
CC	-1 SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC	THIS IS A B TYPE LAMIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL; X54098; CAA38032.1; .
DR	PIR; S21609; S21609.
DR	PIR; B48315; B48315.
DR	MGI; MGI:96796; Lmbd2.
DR	InterPro; IPRO01664; IF.
DR	InterPro; IPRO01322; IF_tall.
DR	InterPro; IPRO01230; Prenyl_sste.
DR	Pfam; PF00038; Filament_1.
DR	Pfam; PF00932; IF_tail; 1.
DR	PROSITE; PS00226; IF; 1.
KW	Intermediate filament; Coiled coil; Nuclear protein; Lipoprotein;
KW	Prexylation; Phosphorylation; Alternative splicing.
FT	DOMAIN 1
FT	FT DOMAIN 27 378
FT	HEAD.
FT	FT DOMAIN 379 592
FT	TAIL.
FT	FT DOMAIN 27 61
FT	COIL 1A.
FT	FT DOMAIN 62 73
FT	LINNER 1.
FT	FT DOMAIN 74 207
FT	COIL 1B.
FT	FT DOMAIN 208 234
FT	LINNER 2.
FT	FT DOMAIN 235 378
FT	COIL 2
FT	FT DOMAIN 414 419
FT	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	ASP/GLU-RICH (HIGHLY ACIDIC, COULD BE
FT	INVOLVED IN CHROMATIN BINDING).
FT	FT DOMAIN 561 574
FT	FARNESYL (BY SIMILARITY).
SO	SEQUENCE 589 589
SO	SEQUENCE 592 AA; 67029 MW; 7D5AC51BC9A5041E CRC64;
Alignment Scores:	
Pred. No.:	4.39
Score:	103.00
Length: 592	
Matches: 63	

	Percent Similarity:	41.44%	Conservative:	58
	Best Local Similarity:	21.58%	Mismatches:	105
	Query Match:	4.21%	Indels:	66
DB:		1	Gaps:	13
US-09-939-293-1 (1-1358) x ILM2_MOUSE (1-592)				
QY	194	ATTGCACGAAATCATCAGACCCTCATTCCTTGTAGTAGACATTTGATGAGAGACGACTG	255	
DB	162	LeuAlaLysLaIagLuAspGLyHisAlaValAlaLalysLySgLnleuGluLySgluThrLeu	181	
QY	254	TCTTTGGTAACAAGATAGATGACCTTCACTTTCTCTCTCTCAGACCATATGCGTTGATTGAA	313	
DB	162	MetArgValAap-----LeuGlusnAacGcySgLnserLeuGlnGlu	195	
QY	314	GCTATTACTGAATATATAAGCGCTGGTTAT-----ACCTTAACCTTCTTTAACCCAAA	367	
DB	196	GlueLuAla---PheSerLySerValPheGluGluGluValAargLuThrArgArgArg	214	
QY	368	TATACAACTTACTCTGGGAAAAATGAAATGATACAGAGAGAAGATGAAGTGTCGACGTATC	427	
DB	215	HISGLuArgArgLeuValGluValAspSerSerArgInGlnGlnLuTrypArpPheLyMet	234	
QY	428	ATAGACGCCACAGACTGCAGATACATTCAAACACCAACAGATACCTTGAAG-----	475	
DB	235	AlaGlnAlaLeuGluAspLeuArgSerSerIinhIsaspGluGlnValAArgLeuTryArgVal	254	
QY	476	--CTGGAACACCACTTGATGATACATGCAGTGTCGTTTCACAGATGACAGACAGAACTGCA	532	
DB	255	GlueGluGlnGlnThrTYrChlAlaLalysLeuAspspnAlaLylsLeuLeuSerAspGln---	273	
QY	533	TATCAAACTGGCGGCATCAGCGCTCTTTAAACCCACAGAAATCATCATTTCAAGCTGTGAAA	592	
DB	274	-----AsuAspLyAlaAlaAlaHisAlaIaIaArgGluGluLeuLySgluAlaArg	289	
QY	593	CTGCAGGTGGGAAGAGGCT-----CACAGGCTGTCGGGAAAGACAGAAACCAACAGCTGCA	646	
DB	290	MetArgValGluSerLeuSerTyrlGlnLeuLeuGlnLySglnLySglnAlaSerAlaIa	309	
QY	647	GAAGCACAGATAGAGAGACTCCGTCAG-----	673	
DB	310	GIUAsnHISILeHISgLnGluGlnLuAlaLeuArgGlyLuAArgAspLyPheArgLyS	329	
QY	674	-----AAACACAGAGAGAGAGGAGAGCGGCGCTGACTGCGACAGAG--	718	
DB	330	MetLeuAspAlaLylsGlnGlnGlnMetThrGluValAArgAspArgMetGlnGlnGlnLeu	349	
QY	719	-----GCCTAC	724	
DB	350	AlaGluTYrGlnGlnLeuLeuAspIleLySleuAlaLeuAspMetGluIleSerAlaTyrr	369	
QY	725	CTGCCTGAGATTGAGG--CTTGAGACACACTCCCTGCTCCCCAC-----	768	
DB	370	ArgLySLeuLeuGlnGlyGlnGlnLuArgLeuLySleuSerProSerProSerArg	389	
QY	770	--TGAGGGGAAAGCAGGGGACAGATCCACCTCCGACAGCGTGGCATGACTGTGT	826	
DB	390	IleThrIleSerArgAlaThrsSerSerSerSerSerGlyValGlyMetSerVal---	408	
QY	827	GCACGAGAGAAGCGGCG--AGTCCCTGCCCCTGGCCATTCAGGCGGAGACGCTTTG	880	
DB	409	--GlyGlnArgGlyGlyLysArgArgArgLeuGlnAspThrSerGly--SerProSer	426	
QY	881	TGAGCTGTGATGCTCCTCTGTGGTCTACGCTTGCG	916	
DB	427	ArgAlaSerArgValSerSerGlySerArgLeuAla	438	
RESULT 14				
ID	REST_CHICK	STANDARD:	PRT:	1433 AA.
AC	042184; 042228; 057563; 057564;			
DT	15-JUL-1998 (Rel. 36, Created)			
DR	15-JUL-1998 (Rel. 36, Last sequence update)			

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.,
"440-kD ankyrinB: structure of the major developmentally regulated
RT domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
Lux S.E., Ward D.C., Forget B.G.,
RT "Isolation and chromosomal localization of a novel nonerythroid
RL ankyrin gene.";
Genomics 10:858-866(1991).
CC -1- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -1- PMR: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC AND FUNCTION (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isp-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X56957; CAA40278.1; -;
DR EMBL: X56958; CAA40279.2; -;
DR EMBL: Z26634; CAB42644.1; -;
DR EMBL: M37123; AAA62828.1; -;
DR PIR: S14533; S14533.
DR PIR: A39643; A39643.
DR PIR: B39643; B39643.
DR PIR: S14569; S14569.
DR HSSP: P42771; IDC2.
DR Genew: HGNC:493; ANK2.
DR MIM: 106410; -;
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZUS.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZUS; 1.
DR PRINTS: PRO1415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 20.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation.
FT REPEAT 63 92 ANK 1.
FT REPEAT 96 125 ANK 2.
FT REPEAT 129 158 ANK 3.
FT REPEAT 162 191 ANK 4.
FT REPEAT 193 220 ANK 5.
FT REPEAT 232 261 ANK 6.
FT REPEAT 265 294 ANK 7.
FT REPEAT 298 327 ANK 8.
FT REPEAT 331 360 ANK 9.
FT REPEAT 364 393 ANK 10.

FT REPEAT 397 426 ANK 11.
FT REPEAT 430 459 ANK 12.
FT REPEAT 463 492 ANK 13.
FT REPEAT 496 525 ANK 14.
FT REPEAT 529 558 ANK 15.
FT REPEAT 562 591 ANK 16.
FT REPEAT 595 624 ANK 17.
FT REPEAT 628 657 ANK 18.
FT REPEAT 661 690 ANK 19.
FT REPEAT 694 723 ANK 20.
FT REPEAT 727 756 ANK 21.
FT REPEAT 760 789 ANK 22.
FT REPEAT 793 822 ANK 23.
FT REPEAT 822 850 ANK 24.
FT REPEAT 850 882 ANK 25.
FT REPEAT 882 914 ANK 26.
FT REPEAT 914 946 ANK 27.
FT REPEAT 946 978 ANK 28.
FT REPEAT 978 1010 ANK 29.
FT REPEAT 1010 1042 ANK 30.
FT REPEAT 1042 1074 ANK 31.
FT REPEAT 1074 1106 ANK 32.
FT REPEAT 1106 1138 ANK 33.
FT REPEAT 1138 1170 ANK 34.
FT REPEAT 1170 1202 ANK 35.
FT REPEAT 1202 1234 ANK 36.
FT REPEAT 1234 1266 ANK 37.
FT REPEAT 1266 1298 ANK 38.
FT REPEAT 1298 1330 ANK 39.
FT REPEAT 1330 1362 ANK 40.
FT REPEAT 1362 1394 ANK 41.
FT REPEAT 1394 1426 ANK 42.
FT REPEAT 1426 1458 ANK 43.
FT REPEAT 1458 1490 ANK 44.
FT REPEAT 1490 1522 ANK 45.
FT REPEAT 1522 1554 ANK 46.
FT REPEAT 1554 1586 ANK 47.
FT REPEAT 1586 1618 ANK 48.
FT REPEAT 1618 1650 ANK 49.
FT REPEAT 1650 1682 ANK 50.
FT REPEAT 1682 1714 ANK 51.
FT REPEAT 1714 1746 ANK 52.
FT REPEAT 1746 1778 ANK 53.
FT REPEAT 1778 1810 ANK 54.
FT REPEAT 1810 1842 ANK 55.
FT REPEAT 1842 1874 ANK 56.
FT REPEAT 1874 1906 ANK 57.
FT REPEAT 1906 1938 ANK 58.
FT REPEAT 1938 1970 ANK 59.
FT REPEAT 1970 2002 ANK 60.
FT REPEAT 2002 2034 ANK 61.
FT REPEAT 2034 2066 ANK 62.
FT REPEAT 2066 2098 ANK 63.
FT REPEAT 2098 2130 ANK 64.
FT REPEAT 2130 2162 ANK 65.
FT REPEAT 2162 2194 ANK 66.
FT REPEAT 2194 2226 ANK 67.
FT REPEAT 2226 2258 ANK 68.
FT REPEAT 2258 2290 ANK 69.
FT REPEAT 2290 2322 ANK 70.
FT REPEAT 2322 2354 ANK 71.
FT REPEAT 2354 2386 ANK 72.
FT REPEAT 2386 2418 ANK 73.
FT REPEAT 2418 2450 ANK 74.
FT REPEAT 2450 2482 ANK 75.
FT REPEAT 2482 2514 ANK 76.
FT REPEAT 2514 2546 ANK 77.
FT REPEAT 2546 2578 ANK 78.
FT REPEAT 2578 2610 ANK 79.
FT REPEAT 2610 2642 ANK 80.
FT REPEAT 2642 2674 ANK 81.
FT REPEAT 2674 2706 ANK 82.
FT REPEAT 2706 2738 ANK 83.
FT REPEAT 2738 2770 ANK 84.
FT REPEAT 2770 2802 ANK 85.
FT REPEAT 2802 2834 ANK 86.
FT REPEAT 2834 2866 ANK 87.
FT REPEAT 2866 2898 ANK 88.
FT REPEAT 2898 2930 ANK 89.
FT REPEAT 2930 2962 ANK 90.
FT REPEAT 2962 2994 ANK 91.
FT REPEAT 2994 3026 ANK 92.
FT REPEAT 3026 3058 ANK 93.
FT REPEAT 3058 3090 ANK 94.
FT REPEAT 3090 3122 ANK 95.
FT REPEAT 3122 3154 ANK 96.
FT REPEAT 3154 3186 ANK 97.
FT REPEAT 3186 3218 ANK 98.
FT REPEAT 3218 3250 ANK 99.
FT REPEAT 3250 3282 ANK 100.
FT REPEAT 3282 3314 ANK 101.
FT REPEAT 3314 3346 ANK 102.
FT REPEAT 3346 3378 ANK 103.
FT REPEAT 3378 3410 ANK 104.
FT REPEAT 3410 3442 ANK 105.
FT REPEAT 3442 3474 ANK 106.
FT REPEAT 3474 3506 ANK 107.
FT REPEAT 3506 3538 ANK 108.
FT REPEAT 3538 3570 ANK 109.
FT REPEAT 3570 3602 ANK 110.
FT REPEAT 3602 3634 ANK 111.
FT REPEAT 3634 3666 ANK 112.
FT REPEAT 3666 3698 ANK 113.
FT REPEAT 3698 3730 ANK 114.
FT REPEAT 3730 3762 ANK 115.
FT REPEAT 3762 3794 ANK 116.
FT REPEAT 3794 3826 ANK 117.
FT REPEAT 3826 3858 ANK 118.
FT REPEAT 3858 3890 ANK 119.
FT REPEAT 3890 3922 ANK 120.
FT REPEAT 3922 3954 ANK 121.
FT REPEAT 3954 3986 ANK 122.
FT REPEAT 3986 4018 ANK 123.
FT REPEAT 4018 4050 ANK 124.
FT REPEAT 4050 4082 ANK 125.
FT REPEAT 4082 4114 ANK 126.
FT REPEAT 4114 4146 ANK 127.
FT REPEAT 4146 4178 ANK 128.
FT REPEAT 4178 4210 ANK 129.
FT REPEAT 4210 4242 ANK 130.
FT REPEAT 4242 4274 ANK 131.
FT REPEAT 4274 4306 ANK 132.
FT REPEAT 4306 4338 ANK 133.
FT REPEAT 4338 4370 ANK 134.
FT REPEAT 4370 4402 ANK 135.
FT REPEAT 4402 4434 ANK 136.
FT REPEAT 4434 4466 ANK 137.
FT REPEAT 4466 4498 ANK 138.
FT REPEAT 4498 4530 ANK 139.
FT REPEAT 4530 4562 ANK 140.
FT REPEAT 4562 4594 ANK 141.
FT REPEAT 4594 4626 ANK 142.
FT REPEAT 4626 4658 ANK 143.
FT REPEAT 4658 4690 ANK 144.
FT REPEAT 4690 4722 ANK 145.
FT REPEAT 4722 4754 ANK 146.
FT REPEAT 4754 4786 ANK 147.
FT REPEAT 4786 4818 ANK 148.
FT REPEAT 4818 4850 ANK 149.
FT REPEAT 4850 4882 ANK 150.
FT REPEAT 4882 4914 ANK 151.
FT REPEAT 4914 4946 ANK 152.
FT REPEAT 4946 4978 ANK 153.
FT REPEAT 4978 5010 ANK 154.
FT REPEAT 5010 5042 ANK 155.
FT REPEAT 5042 5074 ANK 156.
FT REPEAT 5074 5106 ANK 157.
FT REPEAT 5106 5138 ANK 158.
FT REPEAT 5138 5170 ANK 159.
FT REPEAT 5170 5202 ANK 160.
FT REPEAT 5202 5234 ANK 161.
FT REPEAT 5234 5266 ANK 162.
FT REPEAT 5266 5298 ANK 163.
FT REPEAT 5298 5330 ANK 164.
FT REPEAT 5330 5362 ANK 165.
FT REPEAT 5362 5394 ANK 166.
FT REPEAT 5394 5426 ANK 167.
FT REPEAT 5426 5458 ANK 168.
FT REPEAT 5458 5490 ANK 169.
FT REPEAT 5490 5522 ANK 170.
FT REPEAT 5522 5554 ANK 171.
FT REPEAT 5554 5586 ANK 172.
FT REPEAT 5586 5618 ANK 173.
FT REPEAT 5618 5650 ANK 174.
FT REPEAT 5650 5682 ANK 175.
FT REPEAT 5682 5714 ANK 176.
FT REPEAT 5714 5746 ANK 177.
FT REPEAT 5746 5778 ANK 178.
FT REPEAT 5778 5810 ANK 179.
FT REPEAT 5810 5842 ANK 180.
FT REPEAT 5842 5874 ANK 181.
FT REPEAT 5874 5906 ANK 182.
FT REPEAT 5906 5938 ANK 183.
FT REPEAT 5938 5970 ANK 184.
FT REPEAT 5970 6002 ANK 185.
FT REPEAT 6002 6034 ANK 186.
FT REPEAT 6034 6066 ANK 187.
FT REPEAT 6066 6098 ANK 188.
FT REPEAT 6098 6130 ANK 189.
FT REPEAT 6130 6162 ANK 190.
FT REPEAT 6162 6194 ANK 191.
FT REPEAT 6194 6226 ANK 192.
FT REPEAT 6226 6258 ANK 193.
FT REPEAT 6258 6290 ANK 194.
FT REPEAT 6290 6322 ANK 195.
FT REPEAT 6322 6354 ANK 196.
FT REPEAT 6354 6386 ANK 197.
FT REPEAT 6386 6418 ANK 198.
FT REPEAT 6418 6450 ANK 199.
FT REPEAT 6450 6482 ANK 200.
FT REPEAT 6482 6514 ANK 201.
FT REPEAT 6514 6546 ANK 202.
FT REPEAT 6546 6578 ANK 203.
FT REPEAT 6578 6610 ANK 204.
FT REPEAT 6610 6642 ANK 205.
FT REPEAT 6642 6674 ANK 206.
FT REPEAT 6674 6706 ANK 207.
FT REPEAT 6706 6738 ANK 208.
FT REPEAT 6738 6770 ANK 209.
FT REPEAT 6770 6802 ANK 210.
FT REPEAT 6802 6834 ANK 211.
FT REPEAT 6834 6866 ANK 212.
FT REPEAT 6866 6898 ANK 213.
FT REPEAT 6898 6930 ANK 214.
FT REPEAT 6930 6962 ANK 215.
FT REPEAT 6962 6994 ANK 216.
FT REPEAT 6994 7026 ANK 217.
FT REPEAT 7026 7058 ANK 218.
FT REPEAT 7058 7090 ANK 219.
FT REPEAT 7090 7122 ANK 220.
FT REPEAT 7122 7154 ANK 221.
FT REPEAT 7154 7186 ANK 222.
FT REPEAT 7186 7218 ANK 223.
FT REPEAT 7218 7250 ANK 224.
FT REPEAT 7250 7282 ANK 225.
FT REPEAT 7282 7314 ANK 226.
FT REPEAT 7314 7346 ANK 227.
FT REPEAT 7346 7378 ANK 228.
FT REPEAT 7378 7410 ANK 229.
FT REPEAT 7410 7442 ANK 230.
FT REPEAT 7442 7474 ANK 231.
FT REPEAT 7474 7506 ANK 232.
FT REPEAT 7506 7538 ANK 233.
FT REPEAT 7538 7570 ANK 234.
FT REPEAT 7570 7602 ANK 235.
FT REPEAT 7602 7634 ANK 236.
FT REPEAT 7634 7666 ANK 237.
FT REPEAT 7666 7698 ANK 238.
FT REPEAT 7698 7730 ANK 239.
FT REPEAT 7730 7762 ANK 240.
FT REPEAT 7762 7794 ANK 241.
FT REPEAT 7794 7826 ANK 242.
FT REPEAT 7826 7858 ANK 243.
FT REPEAT 7858 7890 ANK 244.
FT REPEAT 7890 7922 ANK 245.
FT REPEAT 7922 7954 ANK 246.
FT REPEAT 7954 7986 ANK 247.
FT REPEAT 7986 8018 ANK 248.
FT REPEAT 8018 8050 ANK 249.
FT REPEAT 8050 8082 ANK 250.
FT REPEAT 8082 8114 ANK 251.
FT REPEAT 8114 8146 ANK 252.
FT REPEAT 8146 8178 ANK 253.
FT REPEAT 8178 8210 ANK 254.
FT REPEAT 8210 8242 ANK 255.
FT REPEAT 8242 8274 ANK 256.
FT REPEAT 8274 8306 ANK 257.
FT REPEAT 8306 8338 ANK 258.
FT REPEAT 8338 8370 ANK 259.
FT REPEAT 8370 8402 ANK 260.
FT REPEAT 8402 8434 ANK 261.
FT REPEAT 8434 8466 ANK 262.
FT REPEAT 8466 8498 ANK 263.
FT REPEAT 8498 8530 ANK 264.
FT REPEAT 8530 8562 ANK 265.
FT REPEAT 8562 8594 ANK 266.
FT REPEAT 8594 8626 ANK 267.
FT REPEAT 8626 8658 ANK 268.
FT REPEAT 8658 8690 ANK 269.
FT REPEAT 8690 8722 ANK 270.
FT REPEAT 8722 8754 ANK 271.
FT REPEAT 8754 8786 ANK 272.
FT REPEAT 8786 8818 ANK 273.
FT REPEAT 8818 8850 ANK 274.
FT REPEAT 8850 8882 ANK 275.
FT REPEAT 8882 8914 ANK 276.
FT REPEAT 8914 8946 ANK 277.
FT REPEAT 8946 8978 ANK 278.
FT REPEAT 8978 9010 ANK 279.
FT REPEAT 9010 9042 ANK 280.
FT REPEAT 9042 9074 ANK 281.
FT REPEAT 9074 9106 ANK 282.
FT REPEAT 9106 9138 ANK 283.
FT REPEAT 9138 9170 ANK 284.
FT REPEAT 9170 9202 ANK 285.
FT REPEAT 9202 9234 ANK 286.
FT REPEAT 9234 9266 ANK 287.
FT REPEAT 9266 9298 ANK 288.
FT REPEAT 9298 9330 ANK 289.
FT REPEAT 9330 9362 ANK 290.
FT REPEAT 9362 9394 ANK 291.
FT REPEAT 9394 9426 ANK 292.
FT REPEAT 9426 9458 ANK 293.
FT REPEAT 9458 9490 ANK 294.
FT REPEAT 9490 9522 ANK 295.
FT REPEAT 9522 9554 ANK 296.
FT REPEAT 9554 9586 ANK 297.
FT REPEAT 9586 9618 ANK 298.
FT REPEAT 9618 9650 ANK 299.
FT REPEAT 9650 9682 ANK 300.
FT REPEAT 9682 9714 ANK 301.
FT REPEAT 9714 9746 ANK 302.
FT REPEAT 9746 9778 ANK 303.
FT REPEAT 9778 9810 ANK 304.
FT REPEAT 9810 9842 ANK 305.
FT REPEAT 9842 9874 ANK 306.
FT REPEAT 9874 9906 ANK 307.
FT REPEAT 9906 9938 ANK 308.
FT REPEAT 9938 9970 ANK 309.
FT REPEAT 9970 10002 ANK 310.
FT REPEAT 10002 10034 ANK 311.
FT REPEAT 10034 10066 ANK 312.
FT REPEAT 10066 10098 ANK 313.
FT REPEAT 10098 10130 ANK 314.
FT REPEAT 10130 10162 ANK 315.
FT REPEAT 10162 10194 ANK 316.
FT REPEAT 10194 10226 ANK 317.
FT REPEAT 10226 10258 ANK 318.
FT REPEAT 10258 10290 ANK 319.
FT REPEAT 10290 10322 ANK 320.
FT REPEAT 10322 10354 ANK 321.
FT REPEAT 10354 10386 ANK 322.
FT REPEAT 10386 10418 ANK 323.
FT REPEAT 10418 10450 ANK 324.
FT REPEAT 10450 10482 ANK 325.
FT REPEAT 10482 10514 ANK 326.
FT REPEAT 10514 10546 ANK 327.
FT REPEAT 10546 10578 ANK 328.
FT REPEAT 10578 10610 ANK 329.
FT REPEAT 10610 10642 ANK 330.
FT REPEAT 10642 10674 ANK 331.
FT REPEAT 10674 10706 ANK 332.
FT REPEAT 10706 10738 ANK 333.
FT REPEAT 10738 10770 ANK 334.
FT REPEAT 10770 10802 ANK 335.
FT REPEAT 10802 10834 ANK 336.
FT REPEAT 10834 10866 ANK 337.
FT REPEAT 10866 10898 ANK 338.
FT REPEAT 10898 10930 ANK 339.
FT REPEAT 10930 10962 ANK 340.
FT REPEAT 10962 10994 ANK 341.
FT REPEAT 10994 11026 ANK 342.
FT REPEAT 11026 11058 ANK 343.
FT REPEAT 11058 11090 ANK 344.
FT REPEAT 11090 11122 ANK 345.
FT REPEAT 11122 11154 ANK 346.
FT REPEAT 11154 11186 ANK 347.
FT REPEAT 11186 11218 ANK 348.
FT REPEAT 11218 11250 ANK 349.
FT REPEAT 11250 11282 ANK 350.
FT REPEAT 11282 11314 ANK 351.
FT REPEAT 11314 11346 ANK 352.
FT REPEAT 11346 11378 ANK 353.
FT REPEAT 11378 11410 ANK 354.
FT REPEAT 11410 11442 ANK 355.
FT REPEAT 11442 11474 ANK 356.
FT REPEAT 11474 11506 ANK 357.
FT REPEAT 11506 11538 ANK 358.
FT REPEAT 11538 11570 ANK 359.
FT REPEAT 11570 11602 ANK 360.
FT REPEAT 11602 11634 ANK 361.
FT REPEAT 11634 11666 ANK 362.
FT REPEAT 11666 11698 ANK 363.
FT REPEAT 11698 11730 ANK 364.
FT REPEAT 11730 11762 ANK 365.
FT REPEAT 11762 11794 ANK 366.
FT REPEAT 11794 11826 ANK 367.
FT REPEAT 11826 11858 ANK 368.
FT REPEAT 11858 11890 ANK 369.
FT REPEAT 11890 11922 ANK 370.
FT REPEAT 11922 11954 ANK 371.
FT REPEAT 11954 11986 ANK 372.
FT REPEAT 11986 12018 ANK 373.
FT REPEAT 12018 12050 ANK 374.
FT REPEAT 12050 12082 ANK 375.
FT REPEAT 12082 12114 ANK 376.
FT REPEAT 12114 12146 ANK 377.
FT REPEAT 12146 12178 ANK 378.
FT REPEAT 12178 12210 ANK 379.
FT REPEAT 12210 12242 ANK 380.
FT REPEAT 12242 12274 ANK 381.
FT REPEAT 12274 12306 ANK 382.
FT REPEAT 12306 12338 ANK 383.
FT REPEAT 12338 12370 ANK 384.
FT REPEAT 12370 12402 ANK 385.
FT REPEAT 12402 12434 ANK 386.
FT REPEAT 12434 12466 ANK 387.
FT REPEAT 12466 12498 ANK 388.
FT REPEAT 12498 12530 ANK 389.
FT REPEAT 12530 12562 ANK 390.
FT REPEAT 12562 12594 ANK 391.
FT REPEAT 12594 12626 ANK 392.
FT REPEAT 12626 12658 ANK 393.
FT REPEAT 12658 12690 ANK 394.
FT REPEAT 12690 12722 ANK 395.
FT REPEAT 12722 12754 ANK 396.
FT REPEAT 12754 12786 ANK 397.
FT REPEAT 12786 12818 ANK 398.
FT REPEAT 12818 12850 ANK 399.
FT REPEAT 12850 12882 ANK 400.
FT REPEAT 12882 12914 ANK 401.
FT REPEAT 12914 12946 ANK 402.
FT REPEAT 12946 12978 ANK 403.
FT REPEAT 12978 13010 ANK 404.
FT REPEAT 13010 13042 ANK 405.
FT REPEAT 13042 13074 ANK 406.
FT REPEAT 13074 13106 ANK 407.
FT REPEAT 13106 13138 ANK 408.
FT REPEAT 13138 13170 ANK 409.
FT REPEAT 13170 13202 ANK 410.
FT REPEAT 13202 13234 ANK 411.
FT REPEAT 13234 13266 ANK 412.
FT REPEAT 13266 13298 ANK 413.
FT REPEAT 13298 13330 ANK 414.
FT REPEAT 13330 13362 ANK 415.
FT REPEAT 13362 13394 ANK 416.
FT REPEAT 13394 13426 ANK 417.
FT REPEAT 13426 13458 ANK 418.
FT REPEAT 13458 13490 ANK 419.
FT REPEAT 13490 13522 ANK 420.
FT REPEAT 13522 13554 ANK 421.
FT REPEAT 13554 13586 ANK 422.
FT REPEAT 13586 13618 ANK 423.
FT REPEAT 13618 13650 ANK 424.
FT REPEAT 13650 13682 ANK 425.
FT REPEAT 13682 13714 ANK 426.
FT REPEAT 13714 13746 ANK 427.
FT REPEAT 13746 13778 ANK 428.
FT REPEAT 13778 13810 ANK 429.
FT REPEAT 13810 13842 ANK 430.
FT REPEAT 13842 13874 ANK 431.
FT REPEAT 13874 13906 ANK 432.
FT REPEAT 13906 13938 ANK 433.
FT REPEAT 13938 13970 ANK 434.
FT REPEAT 13970 14002 ANK 435.
FT REPEAT 14002 14034 ANK 436.
FT REPEAT 14034 14066 ANK 437.
FT REPEAT 14066 14098 ANK 438.
FT REPEAT 14098 14130 ANK 439.
FT REPEAT 14130 14162 ANK 440.
FT REPEAT 14162 14194 ANK 441.
FT REPEAT 14194 14226 ANK 442.
FT REPEAT 14226 14258 ANK 443.
FT REPEAT 14258 14290 ANK 444.
FT REPEAT 14290 14322 ANK 445.
FT REPEAT 14322 14354 ANK 446.
FT REPEAT 14354 14386 ANK 447.
FT REPEAT 14386 14418 ANK 448.
FT REPEAT 14418 14450 ANK 449.
FT REPEAT 14450 14482 ANK 450.
FT REPEAT 14482 14514 ANK 451.
FT REPEAT 14514 14546 ANK 452.
FT REPEAT 14546 14578 ANK 453.
FT REPEAT 14578 14610 ANK 454.
FT REPEAT 14610 14642 ANK 455.
FT REPEAT 14642 14674 ANK 456.
FT REPEAT 14674 14706 ANK 457.
FT REPEAT 14706 14738 ANK 458.
FT REPEAT 14738 14770 ANK 459.
FT REPEAT 14770 14802 ANK 460.
FT REPEAT 14802 14834 ANK 461.
FT REPEAT 14834 14866 ANK 462.
FT REPEAT 14866 14898 ANK 463.
FT REPEAT 14898 14930 ANK 464.
FT REPEAT 14930 14962 ANK 465.
FT REPEAT 14962 14994 ANK 466.
FT REPEAT 14994 15026 ANK 467.
FT REPEAT 15026 15058 ANK 468.
FT REPEAT 15058 15090 ANK 469.
FT REPEAT 15090 15122 ANK 470.
FT REPEAT 15122 15154 ANK 471.
FT REPEAT 15154 15186 ANK 472.
FT REPEAT 15186 15218 ANK 473.
FT REPEAT 15218 15250 ANK 474.
FT REPEAT 15250 15282 ANK 475.
FT REPEAT 15282 15314 ANK 476.
FT REPEAT 15314 15346 ANK 477.
FT REPEAT 15346 15378 ANK 478.
FT REPEAT 15378 15410 ANK 479.
FT REPEAT 15410 15442 ANK 480.
FT REPEAT 15442 15474 ANK 481.
FT REPEAT 15474 15506 ANK 482.
FT REPEAT 15506 15538 ANK 483.
FT REPEAT 15538 15570 ANK 484.
FT REPEAT 15570 15602 ANK 485.
FT REPEAT 15602 15634 ANK 486.
FT REPEAT 15634 15666 ANK 487.
FT REPEAT 15666 15698 ANK 488.
FT REPEAT 15698 15730 ANK 489.
FT REPEAT 15730 15762 ANK 490.
FT REPEAT 15762 15794 ANK 491.
FT REPEAT 15794 1582

```
QY 248 GCAGTGTCTTTG-----GTAACAGATAGCACCCTCT-----ACC 280
      :: ||| ||| |||::|||
Db 1377 ArgLeuProLeuHeuValLysValArgAspThrGlnGluProCysGlyArgLeuSer 1396
      ::|||
QY 281 TTTCTCTCTCAG-----ACCACATATGCGTGTATT--GAAGCTATT-----ACTGAA 325
      |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
Db 1397 PheMetLysGluProLysSerThrArgGlyLeuValHisGlnAlaIleCysAsnLeuAsn 1416
      ::|||
QY 326 TATACTRAGGCTGTTTATTACTTCTCTTACCGACATATACAGTTTACTTGGC 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1417 IleThrLeuProIleThrLysGluSerGluSerAspGlnGluGluIle 1436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 AAATGAAATTCAGAGAGGAAGATGAA----- 412
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1437 AspMetThrSerGluLysAsnAspGluThrGluSerThrGluThrSerValLeuLysSer 1456
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 -----GTGTGGCAGGTGATCATAGAGCCAGAGCTGAGATGACTTCMAAACACAGAG 466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1457 HisLeuValasnGluValProValLeuAlaSerProAspLeuSerGluValSerGlu 1476
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 467 TACTTGAGCTGGAACCACTTGATGACTGCAGTTGGTCTTCAGAGATGGCA----- 520
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1477 ---MetLysGlnAspLeuIleLysMetThrAlaIleLeuThrThrAspValSerAspLys 1495
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 521 -----GCAGAAGCTGCATATCAACTGCGCGCAGATCAG 553
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1496 AlaGlySerIleLysValLysGluLeuValLysAlaAlaGluGluGluProGlyGluPro 1515
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 GCCTCTATTAAACCGCCAGAAATCATCATTCAGCTGGTGAACCTGCAGTGGAAGAGTGCAC 613
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1516 PheGluIleValGluArg-----ValLysGlnAspLeuGluLysValAsn 1530
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 CAGCTCTCCCGGAAGCAAGCAAAACCAAGCTGGCAGACACAGATAGAAGAGCTCCGTCAG 673
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1531 GluIleLeuArgSerGlyThrCysThrArgAspGluSerValGlnSerSerArgSer 1550
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 674 AAA-----ACACAGAGAGAA-----GGGAGAGACGGGCTGAGCTCG 709
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1551 GluArgGlyLeuValGluGluGluTrpValIleValSerAspGluIleGluGluAla 1570
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 710 GAGCAGAGGCGCTACCTGCGCTGAGGATTGAGGGCT--GAGCACACTGCCCTGCTGCC 766
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1571 ArgGlnLysAlaProLeuGluIleThrGluLysProCysValGluValArgIleAspLys 1590
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 767 CACTCAAGTGGGGAAGAGAGGCGCAGATGCCACCTGCCCAAGGTTGGC----- 814
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1591 GluIleLysGlyLysValGluLysAspSerThr-----GlyLeuValAsnTyr 1606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 815 -----ATGACTGTCTGTGCACCG----- 832
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1607 LeuThrAspAspLeuAsnThrCysValProLeuProLysGluGlnLeuGlnThrValGln 1626
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 AGAAGAGCGCGCAGGCTCGCCT-----GCCCAATCAGGCGAG 871
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1627 AspLysAlaGlyLysLysCysGluAlaLeuAlaValGlyArgSerSerGlu 1643
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 20, 2003, 14:58:18
Job time : 47.5 secs

THIS PAGE BLANK (USPTO)

Query Match:	25.06%	Indels:	2
DB:	11	Gaps:	1
US-09-939-293-1 (1-1358) x Q8RI08 (1-157)			
QY 20 ATGCGCGGTCTGAGAGATGGCTTCGCGGACGCTAACTTCAATCTTCAGGTACAGAC	79		
DB 1 MetabalaalabaYserTTPvalThrArgSerValCysSerLeuPheArgTyrArgGln	20		
QY 80 TGTGTGTCTGTCCTGTGTGGCTTAACCTTAAAGACGGGTCTTCTCAGATGATGAGA	139		
DB 21 ArgPhe-----ProValLeuAlaAsnSerLysArgCysPheSerGluLeuIleLys	38		
QY 140 CCATGGCACAACACTGTGACGATTGGCTTTGGAGTAAACCTGTGTGCGGTCTTATGCA	199		
DB 39 ProTrpHisLysThrValLeuThrGlyPheGlyMetThrLeuCysAlaValProIleAla	58		
QY 200 CAGAAATCAGAGCCTCATCTCCCTTAGTATGTAAGACATTGATGAGAGACAGTGTCTTG	255		
DB 59 GlnLysSerGluProGlnSerLeuSerAsnGluAlaLeuMetArgAlaValSerLeu	78		
QY 260 GTACAGATTACACCTCTACTCTTCCTGACGCCACATATGGTGTGATGTAAGCATT	319		
DB 79 ValTrpAspSerThrSerThrPheLeuSerGlnThrThrTyrAlaLeuIleGluAlaIle	98		
QY 320 ACTGATATTAACTAAGCGCTGTATTATACCTTAATCTTCCTTACGACAAATATACAGTTTA	379		
DB 99 ThrGluTyrThrLysAlaValTyrThrLeuValSerLeuTyrArgGlnTyrThrSerLeu	118		
QY 380 CTGGGAAATGAAATTCACAGGACGAGGAAGATGACAGTGGCAGGTGATCATGACGCCAGA	439		
DB 119 LeuGlyLysMetAsnSerGlnGluAspGluValTyrGlnValIleIleGlyAlaArg	138		
QY 440 GCTGAGATGACT 451			
DB 139 ValGluValSer 142			
RESULT 2			
Q9P682 PRELIMINARY; PRT; 803 AA.			
AC Q9P682; 01-OCT-2000 (TREMBlrel. 15, Created)			
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)			
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)			
DE Related to PCF11 component of pre-mRNA 3'-end processing factor CF			
GN B1D1.390.			
OS Neurospora crassa.			
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC Sordariales; Sordariaceae; Neurospora.			
OX NCBI_TaxId=5141;			
RN [1]			
RE SEQUENCE FROM N.A.			
RP Schulte U., Alpn V., Hoheisel J., Brandt P., Fartmann B., Holland R.,			
RA Nyakatura G., Mewes H.W., Mannhaupt G.;			
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
RN [2]			
RE SEQUENCE FROM N.A.			
RP German Neurospora genome project;			
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL: AL355927; CAB91288.2; -			
DR InterPro: IPR002965; P-rich_extensn.			
DR PRINTS: PR01217; PRICHEXTENS.			
DR SEQUENCE 803 AA; 84196 MW; F9955AEC3BC1E5B CRC64;			

Alignment Scores:	
Pred. No.:	0.0769
Score:	123.00
Percent Similarity:	42.41%
Best Local Similarity:	31.01%
Query Match:	5.10%
DB:	3
Length:	8039
Matches:	49
Conservative:	18
Mismatches:	72
Indels:	19
Gaps:	6

[illegible]

Alignment Scores:	
Pred. No.:	0 129
Score:	120.00
Percent Similarity:	39.588
Best Local Similarity:	21.91%
Query Match:	4.91%
DB:	4
	Gaps: 11
	Matches: 600
	Conservative: 52
	Mismatches: 50
	Indels: 87
	Gaps: 84


```

QY 433 TCCTATGAT-----CACCTGCCACACTTCATCTTC----- 404
DB 319 GLYTYRPROGlnAsnTYHisThrSerHisSerTYrPheGLYThrMetGLuAlaSerTYr 338
QY 403 CTCCTGTAATTCATTTTCCCAAGTAACCTGTATATTGTGGTAAGAGAAAGTTAAGGT 344
DB 339 LeuProSerValProPhePro-----GLYAlaAlaProGLYcysGLY 352
QY 343 ATAAACACGCTTAGTATTTGTAATAGCTTCATCAACGCATATGTGCTGTGAGAG 284
DB 353 MetGLyAspMetSerMet-----SerGLYGLnGLnSerAlaSerGLnLeuGLYHis 369
QY 283 AAAGTAGAGGTGCTATCTGTACCAAGACATCTCTCTCATCAATGCTTCACTACT 224
DB 370 HisGLYValGLY-----GLnAlaHisSerAlaProHis----- 380
QY 223 AAGGAATGAGGCTGTGATTTCTGTGCAATAGGAACCCGACACAGGGTTACTCCAAAGCC 164
DB 380 ----- 380
QY 163 AATGCTACAGTTTGTGCATGCTTATTCATATTC 128
DB 381 -----HisGLnPhenMetProGLnSerTYrProTYr 390

RESULT 5
O9KY63 PRELIMINARY; PRT; 606 AA.
AC O9KY63:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SC04801.
GN SC04801 OR SCD63A.12C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL356832; CAB92664.1.
KW Hypothetical protein.
SQ SEQUENCE 606 AA; 62604 MW; 21B609EAFELC5B4A CRC64;

Alignment Scores:
Pred. No.: 0.211 Length: 606
Score: 117.50 Matches: 50
Percent Similarity: 38.46% Conservative: 10
Best Local Similarity: 32.05% Mismatches: 66
Query Match: 4.87% Indels: 30
DB: Gaps: 7

US-09-939-293-1 (1-1358) x O9KY63 (1-606)
QY 924 AGGTCCACGCGAAGCCTGACACAGAGGACACTCACAGCTCACAAAGGCTTCGGCT 865
DB 59 ArgAlaAlaAlaAlaAlaThrGLYProHisProGLnLeuGLYThrGLYProGLYPro 78
QY 864 GAT---TGGCAGGCGAGACCTGCGGCTCTTCGGTGCACACAGCTATGCCAAC 808
DB 79 GLYValTPGGLuGLYLeuValThrSerAlaLeuGLYThrAspArgArgProProPro 98

```

```

QY 807 CTGGGACAG-----GTGGCATTCGCCCTGCTTCCACAGAGGGGAGAC 760
DB 99 GLYGLYGLuArgGLYProGLYAlaAlaAlaAlaLeuLeuAspAlaAlaAlaGLuThr 118
QY 759 AGGGCAGTGTGCTCAGGCGCTCAATCCT-----CACGACAGTAGGCT---CTGTGCTC 709
DB 118 rValArgArgArgAlaGLYLeuArgProAlaArgAlaAlaArgArgProGLuProAlaPr 138
QY 708 GACTCAGCGCCGCTCCTCCCTTCCTCTGTGTCTTCT----- 672
DB 138 OGLuAspProArgProProLeuProProAlaAlaAlaArgArgLeuThrAlaLeuAl 158
QY 671 -----CACGAGCTCTT-----CTATGTGCT 649
DB 158 aaSPArgProGLYAlaAlaGLYGLYArgArgGLYSerAlaProAspLeuMetGLuLe 178
QY 648 TCTGCGACCTTGCTTGTGCTTCCGAGAGAGCTGTCACCTCTTCACCTGCAGTTTC 589
DB 178 uLeuProGLnTrpLeuAlaThrAlaAsnAlaArgGLYPhaAlaAlaProProGLnAlaLe 198
QY 588 ACCAGCTGAATGTGATTTCTGCGGTATAGAGCGCTGATCGGC 543
DB 198 uProAlaLeuLeuAspAlaAlaArg---GLYArgThrAspLeuArg 212

RESULT 6
O21022 PRELIMINARY; PRT; 1133 AA.
AC O21022:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE F59A2.6 protein.
GN F59A2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Pelodderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lighting J.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z34801; CA84332.1;
DR EMBL: Z66514; CA84332.1; JOINED.
DR EMBL: Z66514; CA91344.1;
DR EMBL: Z34801; CA91344.1; JOINED.
DR InterPro: IPR000237; GRIP-domain.
DR Pfam: PF01465; GRIP;
SQ SEQUENCE 1133 AA; 129239 MW; 92015E790A9DE01F CRC64;

Alignment Scores:
Pred. No.: 0.295 Length: 1133
Score: 116.50 Matches: 54
Percent Similarity: 42.73% Conservative: 43
Best Local Similarity: 23.79% Mismatches: 75
Query Match: 4.76% Indels: 55
DB: Gaps: 9

US-09-939-293-1 (1-1358) x O21022 (1-1133)
QY 197 GCACAGAAATCAGAGCCTATTCCTTAGTAGTGAAGCATGTGATGAGAGAGAGCTGTCT 256
DB 436 AlaGLnLysArgSerSerSerGLuGLuLeuGLuThrAlaAsnGLuMetValArgSerLeuThr 455
QY 257 TTGGTAACAGATAGCAGCTTAC----- 280
DB 456 AlaThrLeuGLuAsnSerAsnSerSerGLuThrGLuLeuLysGLnLysLeuGLuThrLeu 475
QY 281 -----TTTCTCTCAGACACATATGCGTGATTTGAGAGCTATTACT--- 322
DB 476 AspLysGLuLeuGLnAlaAlaArgGLnInThrGLuLysAlaLeuThrInGLuGLuLeuVal 495
QY 323 -----GAAATATACTAAGGCTGTTATACCTTA---ACT 352

```

```

Db 496 LeuThrThrSerLeuAlaGluLysGluGlnGlnThrAlaGlnIleGlnAsnLeuGlnThr 515
QY 353 TCTCTTACCCAGCATATATACAGAGTTTACTTGGGAAATGCAATTCAGAGAGAAATGAA 412
Db 516 GlnIleTyrglnMet-----GluValGluLysGluGlnLysValGlu 529
QY 413 GTGTGGCAGGTGATCATATGAGACCAGAGCTGATGACTTCAAAACACCAAGACTTGG 472
Db 530 LeuValLysValGlnLeuGlnGlnAlaAlaGlnSerSerSerAlaGlnLysValLeu 549
QY 473 AAGCTGGAACCCACTTGGATGATGCA-----GTTGGCTTTTCAGAGATGGCAGCAGAA 526
Db 550 ArgAlaGlnIleGlnGlnLeuGlnAlaLysLeuLysAlaValGlnGlnAlaLysAlaGln 569
QY 527 GCTGATATCAATCGCGGAGATCAGGCTTATACCCGCGAGAAATCAC-----577
Db 570 AlaLeuAsnSerLeuLeuAlaGlnLysGlnHisLeuGlnAlaGlnLeuHisGlnLeuGly 589
QY 578 -----ATTTCAGCTGGTGAACCTGCACTGAGAGAGGTGACAGCAGCTC 619
Db 590 ValGluLysGlnGluLysLeuLysLeuIleMetValLysValGlnLeuGlnGlnAlaAlaGlnSer 609
QY 620 TCCCGGAAAGCAGAAACCAAGCTGGCAGAGACAGATAGAGAGCTCCGTCAGAAACA 679
Db 610 SerSerSerValGlnGlnAlaLeu---ArgAlaGlnIleGlnLysLeuGlnAlaLysLeu 628
QY 660 CAGCAGGAAGGGAGGAGGAGCG-----GCTGATGTCGAGCAG 715
Db 629 GlnGlnIleGlnGlnLysLysAsnAlaLeuAsnAlaSerLeuAlaLysGlnGln 648
QY 716 GAG---GGCTACCTGCTGAG 733
Db 649 GlnThrAlaGlnIleGlnGln 655

```

RESULT 7

Q69088 PRELIMINARY: PRT: 3164 AA.

```

AC 069088:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Virion protein.
OS human herpesvirus 1.
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID-10298;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-17;
RX MEDLINE-88274327; PubMed-2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McIad D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
DR EMBL: D10879; BAA01682.1; -.
DR InterPro: IPR005210; Herpes_UL36.
DR Pfam: PF03586; Herpes_UL36; 1.
SQ SEQUENCE 3164 AA; 335817 MW; 2E3335F9525F8C71 CRC64;

```

Alignment Scores:

```

Pred. No.: 0.806 Length: 3164
Score: 112.50 Matches: 86
Percent Similarity: 33.44% Conservative: 21
Best Local Similarity: 26.88% Mismatches: 128
Query Match: 4.66% Indels: 85
Db: 12 Gaps: 17

```

US-09-939-293-1 (1-1358) x Q69088 (1-3164)

QY 1252 CAGCTGCTCCTCAGACAGAGTGGGGCAGATCA-----GAG 1217
 ::::::::::: ||:::::::::::: :::

```

Db 2581 GluLeuAlaSerAspSerSerGlyProAlaProLeuAlaProAspProLeuSer 2600
QY 1216 AACACATCAGAAATATACATACAAAGAAATGCTCAAACTGGAGTCCGCCCTCC 1157
Db 2601 ProThrAlaAspGlnSerValProThrSerGlnCysAlaProArg-ProProGlyProAl 2620
QY 1156 CACAACT-----GGCATCCCAACAGGAGGAGCACTCAATCAAT 1115
Db 2620 aValThrAlaArgGlnAlaAlaArgProGlyValProAlaGlnSerThrArgProAla----- 2638
QY 1114 TTGACGACGTAAATTAAGACTGMAAAACAGTTTAAACAGTTGCTGAACCTTAAGGGCATGAC 1055
Db 2639 -----ProValGlyProArgAspAspPheArg 2647
QY 1054 AAAA---AGAGCTCTCTCTGACCCAGTGAAGGCAAAATGTTTGGGTGATAGGTA 998
Db 2647 GArgLeuProSerProGlnSerSerProGlyProAspAlaThrAlaProArgProArg 2667
QY 997 AAAATGGGTAAAGACAGCTGTACAGAGTGGGGTGAATGTTAAACGGGTGACGTGCCA 938
Db 2667 cAlaSerSerArgAla-----SerAlaAlaSerSerSe 2678
QY 937 AGGCTTAAGAACCCAGTCCAGCCAGCCAGTGAACAGGAGGACATCAGCTCAGCA 878
Db 2678 rGlySerArgAlaArgAlaArgHisArgAlaArgSerLeuAlaArgAlaThrGlnAlaSe 2698
QY 877 AGGCTTCGCTGATTTGGCCAGGAGGAGACTGCCGCTCTTCGCTGCGGCACACAGCT 818
Db 2698 rAlaThrThrGlnGlnLysTrp-----ArgProProAlaLeu---ProAspThrVal 2713
QY 817 CATGCCAACCCCTGGGCGAGGGTG---GCATCTGCCCTGCTTCCCACTGAGTGGGAGGA 761
Db 2713 lAlaProValThrAspPheAlaAlaArgProProAlaProPolysProProGlnProAlaArg 2733
QY 760 CAGGCGAGTGTCTCAGGCGCTC-----AATCTTCACCCAGGTAGCGCTCTG 713
Db 2733 cHisAlaLeuValSerGly-ValProLeuProLeuGlyProGlnAlaAlaGlnAlaAlas 2753
QY 712 CTCGCACATCAGCCGCTCCTCCCTCTCGCTGTTTCTGACGAGCTCTTATCTG 653
Db 2753 eProAlaLeuProIleAspProValProProAlaAlaThrGly-----T 2769
QY 652 TGGCTTCGCGAGCTTGGTTT-----CTGCTTCGGGAGAGCTGGTGCAC 608
Db 2769 hValLeuProGlyGlyLysAsnAlaArgArgProProLeuThrSerGly-----ProAlaArg 2787
QY 607 CTCTTCACCTTCAGCTTTCACCCAGCTGAATGTATCTCGGGGTAAAGAGCGCTGATC 548
Db 2787 roThrProProArgVal---ProValGlyGlyProGlnArgArgLeuThrArgPro---- 2804
QY 547 TGGCGCAGTTGATATGCAGCTTTCGTCGCCATCTCTGAAAGACCACTGCAGTCA---- 492
Db 2805 -----AlaValAlaSerLeuSerIleuSerArgGlnSerLeuP 2817
QY 491 --TCCAAGTGGTTTCAGCTTCACTCAATCTCTGTTGTTT----- 453
Db 2817 roSerProThrAspProAlaAspProThrAlaProValLeuGlyArgAsnProAlaGlnP 2837
QY 452 --AAGTCATCTCAGCTTGGGCTCTATGATCACTGCACACATTCATCTCTCTCT 399
Db 2837 roThrSerSer-----SerProAlaGlyProSerProProPro 2849

```

RESULT 8

Q94GJ0 PRELIMINARY: PRT: 332 AA.

```

AC 094GJ0:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative retrotransposon protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

CC Ehnhardtoidaeae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buehl C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsilirtin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
RA Vanaken S.E., Utechtack T.R., Feldtlyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0031G04 genomic sequence";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC090714; AAK92683.1; -
SQ SEQUENCE 332 AA; 35100 MW; A552B28D31C7D4FC CRC64;

Alignment Scores:	
Pred. No.:	0.596
Score:	111.50
Percent Similarity:	36.79%
Best Local Similarity:	26.42%
Query Match:	4.62%
DB:	10
Length:	332
Matches:	51
Conservative:	20
Mismatches:	59
Indels:	63
Gaps:	6

QY	899	AGAGGCGCTACAGCTACACAAAGGCGTCGCTGATTGGCCAGGCGAGACCTGCCG	840
Db	139	ArgMetHisSerAspLeuHisLysLysAlaSerLysLysAlaLeuHisnAspGlyLysLeuValArg	158
QY	839	CTCTCTCGGTGGCGACAGACAGCTGATGGCAACCCGCGGCGAGGTGGCGCTGCGCCCTGCTT	780
Db	159	IleIleIleAlaLeuLysAspGlu-----Met	166
QY	779	TCCCCATCGAGTGGGGAGACAGGCGAGCTGTGCTCAGGCCCTCAATCTCCACGAGGTAGG	720
Db	167	AlaArgLeuLysLysGluAsnAlaGlnLeuLysGlyLeuProAlaProArgLysValArg	186
QY	720	-----	720
Db	187	IleArgThrThrProArgLysLeuThrThrAlaProValArgIleGlnLeuAlaProArg	206
QY	719	---CCCTCGCTCGGACGACGACCCGCGTCGTCGCCCTCTGCTGCTGTTTCTGACGAGCT	663
Db	207	AsnProProProProAlaValAlaProAlaAlaProAlaProProAla-----	221
QY	662	CTTCTATCTGTGCTTGTCCGACGCTTGCTGCTTTCGCTTTCGGGAGAGCTGGTGACCTCTT	603
Db	222	-----ValProAla-----AlaProVal	227
QY	602	CCACCTGAGTTTCACCGACGCTGATGTGATTTCTGGCGGTTATAGAGCCTGATCTGCCG	543
Db	228	ProProAlaAlaProAlaAlaAlaPro-----SerProAlaLeuAla	240
QY	542	CAGTTTGATTCAGGCTTCTGCTGCCA-----TCTGTGAAGCCAACTGACATCATCC	489
Db	241	LeuValProValProValLeuAlaProAlaSerAlaLeuSerPheAlaProAlaSerAla	260
QY	488	AAGTGTTTCAGGCTTCAGTACTCTTGGTGTTTGAAGTCATCTCAGCTGGCTCGTA	429
Db	261	AlaArgGlyProAlaSerGlyAsnGlyGlyTyrPheSerAlaThr-ProSerGlySerAla	280
QY	428	TGATCACTCGCCACACTTCATCTTCCTCCCTGAA	394
Db	280	gAspHis-SerSerSerSerSerGlyTyrGln	291

GN U42471 OR N-3AP1.
OS *Mus musculus* (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;

Alignment Scores:	
Pred. No.:	0 658
Score:	111.50
Percent Similarity:	42.86%
Best Local Similarity:	36.51%
Query Match:	4.62%
DB:	11
Length:	520
Matches:	46
Conservative:	51
Mismatches:	8
Indels:	21
Gaps:	5

QY	909	CTGAGACCAACAGAGGACCTACAGACTACAAAGGCGCTTCGGCTGATTGGCCAGGGCAG	850
Db	344	LeuArpPro---ProValIleuGlySerAsnIlySerArgSerProProLeuProProVal	362
QY	849	GACCG-----CGCCCTTCCTGGGTGCACAGACATGATCAGCCACACCTGGGCAG	799
Db	363	ProMetGlyIylAlaProProProProProThProArGlyProProPro-ProGlyArGly	382
QY	798	GTGGCATTCGCCCTCGCTTTTCCCACTGAGAGGGGAGACAGGGCAGTGTCCAGGCCCT	739
Db	382	YelYProProProProProProProAlaThhGlyArgSerGlyProProProProle	402
QY	738	CAATCTTCACGAGGTAGGCTCTGCTCGACTACAGCCGCGTCTCCCTCTCTGT	679
Db	402	u---ProGlyAlaGlyIylProProAlaIleProProProProProProProProPr	421
QY	678	GTTTTCACGAGAGCTCTTATCTGCGCTTCGCCAGCTTGTTCTGGCTTCCGGAG	619
Db	421	OPrOCysProGly-----SerGlyPr	428
QY	618	AGCTGGTCACCTCTTCACCTGCA---GTTTCACACAGCTGAATGTATTCTGGCGTT	562
Db	428	AlaIleProProLeuProProProIlnProValSerGlyIylSerProAlaIleProIylG	448
QY	561	ATAGAGCGCTGATC	548
Db	448	LYATGGIYAlaIleu	452


```

QY 434 GCCAGAGCTGAGTACTTCAAAACACCAAGATGACTTGAAGCTGGAACCACTTGATG 493
Db 1222 -----SerHisHisGluGluIleIleGluAlaAlaAsnGluValVal 1235
QY 494 ACTGACATGGCTCTTTCAGAGATGGCAGACAGAGCTGCATATCAAACTGGCCAGATCAG 553
Db 1236 ArgSerValAspValAspGluLeu-----1243
QY 554 GCCTTATACCGCAGCAATTCATTCAGCTGTGAATTCAGAGTGAAGAGGTGCAC 613
Db 1244 -----AlaArgPheLeuLeuAspLysThrGluProGluAspAspGluAglu 1259
QY 614 CAGCTCTCCGCGAAGACAGAAACCAAG-----640
Db 1260 LysLeuLysLysLysMetGluValThrArgAspGluLeuAlaAspAlaLeuLysGlnLys 1279
QY 641 ---CTGGCAGAAAGACACAGATAGAGAGCTCCGTCAGAAAACACAGAGAGAGGAGAG 697
Db 1280 GlyLeuAlaMetAlaArgIleGluAsnLeuLysGlyGlyGluGlyGluGlyGluGlu 1299
QY 698 CGGCTGAGTCGAGCAGAGAGGCTACCTGCGTGAAGAT 736
Db 1300 -----GlnSerSerGlnLysAspLysPheGluGluAsn 1310

RESULT 12
Q924W1 PRELIMINARY; PRT; 356 AA.
ID 0924W1
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALEX protein.
GN ALEX.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95089824; PubMed=7997272;
RA Kehlenbach R.H., Matthey J., Huttner W.B.;
RT "Xlas is a new type of G protein.";
RL Nature 372:804-809(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Klemke M., Kehlenbach R.H., Huttner W.B.;
RT "Two overlapping reading frames in a single exon encode interacting
RT proteins - a novel way of gene usage.";
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X84047; CAC39212.1;
SQ SEQUENCE 356 AA; 37970 MW; 9849ABD0AE524A3D CRC64;

Alignment Scores:
Pred. No.: 1.46 Length: 356
Score: 107.00 Matches: 46
Percent Similarity: 38.73% Conservative: 9
Best Local Similarity: 32.39% Mismatches: 45
Query Match: 4.43% Indels: 42
DB: 11 Gaps: 8

US-09-939-293-1 (1-1358) x Q924W1 (1-356)
QY 946 CAGTGCACCAAGGCTAAGACAGGTCGACGACAGCTGAGACACAGAGGACCTAC 887
Db 72 GlnSerProThrProLys-----ArgSerProGln-ProAlaGlnProLeuPr 87
QY 886 AGCTCACAAGAGCGCTGCTGATGGCCAGGCGAGGACCTG-----844
Db 87 OAArgThrArgSerLeuProGlyGlnProSerProLeuArgSerProLeuProGl 107
QY 843 -----CCGCCTTCTCGTGGTCACAGACAGTCA 815

```

```

Db 107 YLeuSerLeuLeuProGluProIleGlnProGlyLeu-SerLeuGluProGlnArgC 127
QY 814 GCCAACCCCTGGGACAGGTTGGCATCGCCCTGCTTCCCACTG-----AGTG 767
Db 127 YsgInProLeuLeuGly-----GlnProProLeuGluGlnProMetGlnValLeuTrpSerG 146
QY 766 GGGAGACAGGCGAGTGTCTCAGGCGCTCAATCTCAGCAGAGTAGGCTCTGCTCCGA 707
Db 146 YgluProGlyHis---SerArgLeuLeuGlnPro---LeuGlyHisProSerLeuProA 164
QY 706 CTCAGCCCGCTCCCTCCCTCCCTGCTGTTTTCGAGGAGAGCTTCTATCTGCTTC 647
Db 164 IagInGlnLeuProProGluGlnPro-----LeuL 174
QY 646 TGCCAGCTTGGTCTCTGCTTCCGAGAGGCTGTCACCTTTCACCTGACATTTTCAC 587
Db 174 euProAlaGlnSerLeu-----ProAlaGlyGlnProLeuProProGlnAlaGlyP 191
QY 586 CA 585
Db 191 ro 191

RESULT 13
Q99M73 PRELIMINARY; PRT; 603 AA.
ID 099M73
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Type II 65kd keratin.
GN KRT2-16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C3H/HEN;
RC Polier C., Yoshiki A., Fujiwara K., Guenet J.-L., Kusakabe M.;
RT "Fine mapping of Hague (Hag), a mouse hair mutation with 2
RT epialleles.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028607; AAK6233.1;
DR MGD; MGI:96700; Krt2-16.
DR InterPro; IPR001654; IF.
DR InterPro; IPR003054; Keratin_II.
DR Pfam; PF00038; filament_1.
DR PRINTS; PRO1276; TYPE2KERATIN.
DR PROSITE; PS00226; IF; UNKNOWN_1.
SQ SEQUENCE 603 AA; 64987 MW; AFC01BD8D0E780B9 CRC64;

Alignment Scores:
Pred. No.: 1.64 Length: 603
Score: 107.00 Matches: 42
Percent Similarity: 47.70% Conservative: 41
Best Local Similarity: 24.14% Mismatches: 59
Query Match: 4.37% Indels: 32
DB: 11 Gaps: 7

US-09-939-293-1 (1-1358) x Q99M73 (1-603)
QY 251 GTGCTTTGGTAAACAGATAGCAGCTTACCTTCTCTCTCAGACCACTAT---GCGTTG 307
Db 276 ValAlaLeuLysLysAspValAspAlaAlaPheLeuAsnLysSerAspLeuAlaAsn 295
QY 308 ATTGAAGCTAT-----ACTGAATATACTAAGCTGTTAT-----ACC 346
Db 296 ValAspThrLeuIleGlnGluThrGluLeuLysAlaLeuLysGluGluIleGlu 315
QY 347 TTAACCTCTCTTACCGCAGCATATACAGATTACTGGGAAATGAATTCAGAGAGAA 406
Db 316 MetLeuGlnSerHisIleSerGluThrSerValIleValLysMetAspAsnSerArgAsp 335

```

```

QY 407 GATGAAGTGTGGAGGTGATCATAGAGCCAGAGCTGATGACTTCCAAACACAGAG 466
    ::::: |||||::: ::::: |||||
Db 336 LeuAmLeuaspGlyLeile-----AlaGluValLysAlaGlnTYrGlu 351
    ::::: |||||
QY 467 TACTGAAGTGGAAACCACTTGATGACTGCTGTTTTCAGAGATGACAGAGAA 526
    ::::: |||||
Db 352 ValAlaArg-----ArgSerArgAlaAspValGlu 361
    ::::: |||||
QY 527 GCGCATATCAACAGTGGCCGAGATCAGGCGCTTATACCCGCCAGATTCATTCAGCTG 586
    ::::: |||||
Db 362 SerTrpTyrGlnThrLysTYrGluGluMetArgValThrAlaGlnHisCysAspAsn 381
    ::::: |||||
QY 587 GTGAACCTGAGGTGGAGAGGTGACACAGCTCTCCCG-----AAAGCA 631
    ::::: |||||
Db 382 LeuArgSerThrArgAspGluLeuAsnGluLeuThrArgLeuLeuGlnArgLeuLysAla 401
    ::::: |||||
QY 632 GAACACAGCTGGCAGACAGACAGATAGAAAGCTCCGTCAAGAAACACAGAGAGAGG 691
    ::::: |||||
Db 402 GluIleGluHisThrLysAlaGlnCysAlaLysLeuGluAlaValAlaGlu----- 419
    ::::: |||||
QY 692 GAGGAGCGGCTGAGTCGAGAGAGGCGCTTACTGCTGAG 733
    ::::: |||||
Db 420 -----AlaGluGlnGlnGlyGluAlaAlaLeuAsnAsp 430
    ::::: |||||
RESULT 14
092213 PRELIMINARY; PRT; 738 AA.
AC 092213;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 80.3 kDa protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RA Wang Y.Z., Kehlenbach R.H., Huttner W.B.;
RT "The XL-domain of rat xias is encoded by a single exon.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF093569; AAD03033.1;
DR InterPro: IPR002965; P-rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENS.
KW Hypothetical protein.
SQ SEQUENCE 738 AA; 80340 MW; 51EA2B3A7D9D018A CRC64;

Alignment Scores:
Pred. No.: 1 71 Length: 738
Score: 107.00 Matches: 46
Percent Similarity: 38.73% Conservative: 9
Best Local Similarity: 32.39% Mismatches: 45
Query Match: 4.43% Indels: 42
DB: 11 Gaps: 8

US-09-939-293-1 (1-1358) x 092213 (1-738)
QY 946 CAGTCCCAAGGCGCTAAGAACAGGTCCAGAGCTGAGACACAGAGAGCATCTAC 887
    ||| |||
Db 454 GlnSerProThrProLys-----ArgSerProGln-ProArgGlnProLeuPr 469
    ||| |||
QY 886 AGCTCACAAAGCGCTCTCGCTGATTGGCCAGGCGAGACTG----- 844
    ::::: |||||
Db 469 aArgArgArgSerLeuProGlyGlnProProSerProLeuArgSerProLeuProG 489
    ::::: |||||
QY 843 -----CCGCTCTTCTCGGTCCACAGACAGATCAT 815
    ||| |||
Db 489 yLeuSerLeuLeuProGluProLeuGlnProProGlyLeu-SerLeuGlnProGlnArg 509
    ||| |||
QY 814 GCCAACCTGGGAGGCTGACATCTGCCCTGTTCCCACTG-----AGTG 767
    ||| |||
Db 509 yGlnProLeuLeuGly-----GlnProProLeuGlnGlnPrometGlnValLeuTrpSerG 528
    ||| |||

```

```

QY 766 GGGAGACAGGCGAGTGTCTCAGGCGCTCAATCTCAGCAGGTAGGCGCTGCTCGCA 707
    ||||| |||
Db 528 yGluProGlyHis-----SerArgLeuLeuGlnPro-----LeuGlyHisProSerLeuPro 546
    ::::: |||||
QY 706 CTCAGCCCGCTCCCTCCCTCTCTGTGTTTCTGACGAGAGCTTTATCTGCTTC 647
    ||| |||||
Db 546 LacGlnLeuProProGlnGlnPro-----LeuL 556
    ::::: |||||
QY 646 TGGCAGTTTGTTTCTTCTTCCGGGAGAGCTGTGACCTCTTCCACTTCAGATTTCAC 587
    ||||| |||
Db 556 euProAlaGlnSerLeu-----ProAlaGlnProLeuProProGlnAlaGlyP 573
    ::::: |||||
QY 586 CA 585
    ||
Db 573 TO 573

RESULT 15
09ROT4 PRELIMINARY; PRT; 739 AA.
ID 09ROT4;
AC 09ROT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 84.0 kDa protein (Fragment).
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OX Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EI34;
RA MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Billes D.E., Hollingshead S.K.;
RT "The pspC gene of streptococcus pneumoniae encodes a polymorphic
RT protein, pSPC, which elicits cross-reactive antibodies to Pspa and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1;
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR002965; P-rich_extensn.
DR Pfam: PF01473; CW_binding.1;
DR PRINTS; PRO1574; TUBBYPROTEIN.
DR PRINTS; PRO1217; PRICHEXTENS.
DR TIGRfams; TIGR01168; YSIRK_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83961 MW; 7EE2FE676ABF989 CRC64;

Alignment Scores:
Pred. No.: 1 71 Length: 739
Score: 107.00 Matches: 51
Percent Similarity: 40.09% Conservative: 40
Best Local Similarity: 22.47% Mismatches: 70
Query Match: 4.37% Indels: 66
DB: 2 Gaps: 8

US-09-939-293-1 (1-1358) x 09ROT4 (1-739)
QY 203 AAATCAGAGCGCTATTCCTTAGTAGAGCATTTGATGAGAGAGCAGTGTCTTGCTA 262
    ||||| |||
Db 291 LysSerSerAspSerSerValGlyGluGlnThrLeuProSerProSerLeuAsnMetAla 310
    ::::: |||||
QY 263 ACAGATAGACCTCTTACCTTCTCTCAGACACCATATGCGTTGATGAAGCTATTACT 322
    ::::: |||||
Db 311 AsnGlu-----SerGlnThrGluHisArg-----LysAspValAsp 322
    ::::: |||||
QY 323 GAATATACTAAGCGCTTTTACTTACTTCTTTCACGACAA----- 367
    ||||| |||
Db 323 GluTrpLeuLysLysMetLeuSerGluLeuLeuAspArgLysHisThrGlnAsn 342
    ||||| |||
QY 368 -----TATACAAGTTTACTT 382
    ||||| |||

```

```
Db 343 ValAsnLeuAsnIleLysLeuSerAlaIleLysThrLysTyrLeuTyrGluLeuSerVal 362
OY 383 GCGAAATGAATTCAGAGAGAGAAATGATGCGAGGATCATAGAGCCAGAGCT 442
Db 363 LeuLysGlnSerLysLysGlu-----GluLeuThrSerLysThrLysAla 378
OY 443 GAGATGACTTCAAAACACCAAGAGTACTTGAAAGCTGGAACCACTTGATGACTGCAGTT 502
Db 379 GluLeuThrAlaIaPheGluGlnPheLysLysAspThrLeuLysProGluLysLysVal 398
OY 503 GGTCTTCAAGATGCGAGAGAGCTGCATATCAACATGCGCAGATCAAGCCTCTATA 562
Db 399 AlaGluAlaGluLysLysValGluGluAlaLysLysAlaLysAspGlnLysGluGlu 418
OY 563 ACCGCCAGAAATCAC-----ATTACCTGGTGAACCTGCAGGTG----- 577
Db 419 AspArgArgAsnTyrProThrAsnThrTyrLysThrLeuGluLeuGluIleAlaGluSer 438
OY 578 -----ATTCAGCTGGTGAACCTGCAGGTG----- 601
Db 439 AspValLysValLysLysAlaGluLeuGluLysLysGluGluAlaAsnGluSerArg 458
OY 602 -----GAAGAGGTGCACACGCTCTCCGGAAGAGCAAGCAAGCTGCAGAGCA--- 652
Db 459 AsnGluGluLysLysGlnAlaLysGlnLysValGluSerLysLysAlaGluAlaThr 478
OY 653 -----CAGATAGAAAGACTCCGTACAGAAACACAGAGAGAGGAGAGCGGCT 703
Db 479 ArgLeuGluLysLysLysThrAspArgLysLysAlaGluGluAlaLysArgLysAla 498
OY 704 ---GAGTCGAGCAGAGAGGCC 721
Db 499 GluGluSerGluLysLysAla 505
```

Search completed: February 20, 2003, 15:01.17
Job time : 107 secs

by analysts of the total score after the result being printed

US6110691-A.

✓

Human; caspase activator; Smac; apoptosis; cancer; autoimmune disease
neurodegenerative disease; mitochondria.

OS Homo sapiens
xy

US6110691-A

PD 29-AUG-2000.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Query Match:	4.58%	Indels:	42
DB:	1	Gaps:	8

US-09-939-293-1 (1-1358) x S02041 (1-3660)

```

OY      5  TCCGCGCGCTGCACAAAGTGGCGCTCTCAAGAGTTGGCTGTGGCGACCGTAACCTTCATTC 64
DB      1163 SerLeuAtrgysAapLeuSerGIumethIsGIuTrpIlethrcIdnAlaGIuGIuGIuTr 1182
OY      65  TTCAGGTACAGACAGTGTGTGTGTCTCTGTGGCTACCTTTAAGACGGGTGTTTC 124
DB      1183 LeuGIu-----ArgasphheGIuTrfLysThrPro 1192
OY      125 TCAGAAATTGATTAAGACCATGCGACAAAACTGTGACGATGTGGCTTTGGAGTAACCCCTGT 184
DB      1193 GIuGIuLeuGIuInfs----- 1197
OY      185 GCGGTTCCATTATTCACAGAAATCAGACCCATTCCTTCCCTAGTAGTGACGATGATGAGG 244
DB      1198 AlaValGIuGIuLeuLeuysAtrgAlaIalsGIuApsAlamethGIuInfsGIuVal----- 1214
OY      245 AGAGCGAGTGTCTGTGGTAAGATAGCACTCTACCTTTCTCTCTCAGACCATATTCG 304
DB      1215 ---LysValIleLeuLeuIethrApsSerValAsnaAphIleAlaIalsAalProProAla 1233
OY      305 TTGATTGAAGCTATTACTGCAATTAAGCTGTTATACCTTACTTACTTCTCTTACCGA 364
DB      1234 AlaAsnGIuAlaLeuLeuLysIsgIuleuApsValLeu-----IlethrseryrGIuInfs 1251
OY      365 CAATATCAAGTTACTTACTGGAAATGCAATTCAGAGAGAGAGATAGCTGGCAGGCTG 424
DB      1252 LeuGysSerAtrgLeuAanGIuLysCysLysThrLeuGIu-----GIuValTrp----- 1267
OY      425 ATCATATGAGACCAAGACCTAGATGACTCATAAACCACCAAGACTTGTGAAGCTGGAAC 484
DB      1268 -----AlaCysAtrpIsgIuleuLeuSer-----TyrLeuApsAlaGIuAan 1281
OY      485 ACTGGATACAGCGACTTATTAACCCGCAAGAAATCACATTCAGCTGGTGAACATGCCAGTGGAA 604
DB      1282 LysTrpLeuAanGIuValIalIuleuLysLeuLysAlaThrGIu---AsnIleGIuInfsGIu 1300
OY      545 GCAGATCAGCGCTCTATTAACCCGCAAGAAATCACATTCAGCTGGTGAACATGCCAGTGGAA 604
DB      1301 AlaGIuGIuIleSerGIuSerLeuApsSerLeuGIuAtrgLeuethrAtrgIleProGIuAps 1320
OY      605 GAGGTGCACCACTCTCCCGGAAAGCAGAAACCAAGCTGGCAGAAACACAGATGATAAGAG 664
DB      1321 AsnaAtrgAanGIuIleAtrgIuleuAalGIuInrLeuThrApsGIuLysIleLeuApsGIu 1340
OY      665 CTCCTGCATCAAAAACACAGAGAGAGAGGCGGAGCGGCGTCAATCGGAGACGAGCGCTTAC 724
DB      1341 LeuIleAanGIuLysLeuGIuLysPheAanThrAtrgTrpGIuGIuLeuGIuInfsGIuAla 1360
OY      725 CTGCCTGAGGATTTGAGGGCTGAGCAGCACTGCGCTGTCTCCCACTACATGCGGAGAA 781
DB      1361 ValAtrgAtrgGIuLysSerLeuGIuInserIleGIuInserAlaGIuGIuThrApsLys 1379

RESULT 5
178557
N-methyl-D-aspartate receptor chain NMDAR2D-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: 178557; 158158; D45219
R:Monier, H.; Burnashev, N.; Laurie, D.J.; Sakmann, B.; Seeburg, P.H.
A:Title: Developmental and regional expression in the rat brain and functional properties of the NR1 subunit of the NR1/NR2B NMDA receptor channel
A:Reference number: 158158; PMID:94206533; PMID:7512349
A:Accession: 178557
A:Status: preliminary; translated from GB/EML/DBDUB
A:Molecule type: mRNA
A:Residues: 1-1323 <RES>
A:Cross-references: GB:J31612; NID:g469068; PIDN:AAC37647.1; PID:g469069
A:Accession: 158158

```

A:Status: preliminary; translated from GB/EMBL/JDBJ
A:Molecule type: mRNA
A:Residues: 1-66, 'V', 66-1323 <RE2>
A:Cross-references: GB:IJ31611, NID:q469066; PIDN:AAC37646.1; PID:q469067
R: Ishii, T., Morigoshi, K., Sugihara, H., Sakurada, K., Kadohira, H., Yokoi, M., Akai
J. Biol. Chem. 268, 2836-2843, 1993
A:Title: Molecular characterization of the family of the N-methyl-D-aspartate receptor
A:Reference number: A45219, M0ID:J3155102; PMID:8428958
A:Accession: D45219
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1265-1323 <ISH>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:124265)
C:Superfamily: N-methyl-D-aspartate receptor 2D: glutamate receptor homology
F:451-879/Domain: glutamate receptor homology <GR1>

Alignment Scores:

Pred. NO.:	0.569	Length:	1322
Score:	111.00	Matches:	56
Percent Similarity:	31.378	Conservative:	8
Best Local Similarity:	27.458	Mismatches:	73
Query Match:	4.608	Indels:	67
DB:	2	Gaps:	9

US-09-939-293-1 (1-1358) x I78557 (1-1323)

QY	915	GGAGGCTGGAGACCCAGGAGGAGCACTCAGAGCTACACAAAGCGCTTCGGCTATTGGCCA	856
Db	1004	AlaIleValaArgGluGluGlnGluProThrGluProAlaGlyAlaPheProGlyPhePro	1022
QY	855	GGGCAAGACCTCCGCCCTCTTCTGGGTCACAGACAGCACTCATGCGCAACCCGGGACAGGTG	796
Db	1024	SerProProAlaPheProProAlaAlaAlaAlaAlaValaAlaGlyProProLeucysarGlu	1044
QY	795	GGA-----TTCGCCCCCTTCCTCCACAGTAGGG-----	766
Db	1044	AlaPheGluAspGluSerProProAlaProSerArgrTrpProArgSerAspProGluSer	1063
QY	765	-----GGAGACAGGAGCAGTAGTGTGTCAGAGGCCCTCAATCTCCACAGAGTAGGCC	718
Db	1064	GlnProLeuLeuGlnGlyGlyGlyAla-----GlyGlyProSerAlaGlyAlaProThrAl	1081
QY	717	TCTCTCTCCGAGTCAGCCCGCTCTCC-----CTTCTCTCTGTGT	676
Db	1081	AProProProArgrArgrAlaAlaAProProProGlyAlaTyrLeuAspLeuGluProSerP	1101
QY	675	TTTCGAC-----GAGCCTCTTCTATCTGTGCTCTTCTCCACC	640
Db	1101	OSerAspSerGluAspSerGluSerLeuGlnGlyAlaSerLeuGlnGlyLeuGluProTr	1121
QY	639	TTGGTTCTG-----CTTCCGGGAGAGCGGTGCACCC-----	606
Db	1121	PTrPheAlaAspPheProTyrProTyrAlaGluArgLeuGlyProProProGlyArgTyr	1141
QY	605	-----CTTCCACCTGC	595
Db	1141	rTrpSerValAspLeuGlnGlyTrpArgAlaGlySerTrpAspTyrLeuProProAr	1161
QY	594	AGTTTCAACCACT-----GAAATGTCATTCTCTGGCGTTATAGAGCGCTGATCT	547
Db	1161	gGlyGlyProAlaTrpHisCysArgHisCysAlaSerLeuGlnLeuLeuProProProAr	1181
QY	546	GGCGACATTGATATGCACTCTGGCGGCATCTCGAAGAACCAACAGCAATGATCTCA	487
Db	1181	ghis-----LeuSerCysSerHisAspGlyLeuAspGlyGlyTr	1194
QY	486	GTCGTTTCCA	477
Db	1194	PTPrAlaPro	1197
RESULT	6		
	445219		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 20, 2003, 14:47:03 ; Search time 73.5 Seconds

(without alignments)
4923,928 Million cell updates/sec

Title: US-09-939-293-1

Perfect score: 2446

Sequence: 1 ggcgtccgcgcgcgcacacaa.....ggactaacacagaaaaaa 1358

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ n2p model -DEV-xlp
-O/-cgcn2.1/USPQC/spool/US09939293/runat_20022003.111509.14822/app_query.fasta_1.1543
-DB-A Geneseq_101002 -QFMT-fastan -SUFFIX-frag -MIMMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -STARF=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cd1
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-pio -NORR-ext HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09939293 @CGCN.1.1.47 @runat_20022003.111509.14822 -NCPU=6 -ICPU=3
-NO_XLPHYX -NO_WMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1201	49.1	239	21	AA26210	Human caspase acti
2	1201	49.1	239	21	AA26210	Inhibitor of apopt
3	1138	46.5	227	21	AA26210	Human pancreatic c
4	879	35.9	186	22	AA26210	Human protein sequ
5	190	7.8	40	23	AA26210	Inhibitor of apopt
6	186	7.6	39	23	AA26210	Inhibitor of apopt
7	164	6.7	35	23	AA26210	Inhibitor of apopt
8	141	5.8	30	23	AA26210	Inhibitor of apopt
9	125.5	5.2	502	22	AA26210	Human WASP protein
10	125.5	5.2	502	22	AA26210	Amino acid sequenc
11	125.5	5.2	502	22	AA26210	Human WASP protein
12	119.5	5.0	317	22	AA26210	Human WASP protein
13	119.5	5.0	317	22	AA26210	Human WASP protein
14	116	4.8	276	22	AA26210	Human WASP protein
15	112	4.6	1485	21	AA26210	Human WASP protein
16	111.5	4.6	520	22	AA26210	Human WASP protein
17	111.5	4.6	520	22	AA26210	Human WASP protein
18	111	4.5	520	22	AA26210	Human WASP protein
19	110	4.5	589	22	AA26210	Human WASP protein
20	108	4.4	525	23	AA26210	Human WASP protein
21	107.5	4.4	1396	23	AA26210	Human WASP protein
22	107	4.4	589	20	AA26210	Human WASP protein
23	107	4.4	589	20	AA26210	Human WASP protein
24	107	4.4	529	18	AA26210	Human WASP protein
25	107	4.4	529	20	AA26210	Human WASP protein
26	106.5	4.4	370	22	AA26210	Human WASP protein
27	106.5	4.4	370	22	AA26210	Human WASP protein
28	106.5	4.4	370	22	AA26210	Human WASP protein
29	106	4.4	520	22	AA26210	Human WASP protein
30	106	4.4	520	22	AA26210	Human WASP protein
31	104.5	4.3	591	19	AA26210	Human WASP protein
32	104	4.3	663	22	AA26210	Human WASP protein
33	104	4.3	749	22	AA26210	Human WASP protein
34	103.5	4.2	317	22	AA26210	Human WASP protein
35	101.5	4.1	1197	22	AA26210	Human WASP protein
36	101.5	4.1	1297	22	AA26210	Human WASP protein
37	101	4.1	441	22	AA26210	Human WASP protein
38	101	4.1	1072	23	AA26210	Human WASP protein
39	101	4.1	1719	23	AA26210	Human WASP protein
40	100.5	4.1	1084	23	AA26210	Human WASP protein
41	100	4.1	337	22	AA26210	Human WASP protein
42	100	4.1	518	22	AA26210	Human WASP protein
43	100	4.1	551	21	AA26210	Human WASP protein
44	100	4.1	564	19	AA26210	Human WASP protein
45	100	4.1	581	20	AA26210	Human WASP protein

ALIGNMENTS

RESULT 1	ALIGNMENTS
AA26210	AA26210 standard; Protein; 239 AA.
XX	AA26210:
XX	23-FEB-2001 (first entry)
DT	Human caspase activator Smac.
XX	Human caspase activator Smac.
DE	Human caspase activator Smac.
XX	Human caspase activator Smac.
XX	Human caspase activator Smac.
KW	Human: caspase activator; Smac; apoptosis; cancer; autoimmune disease;
KM	neurodegenerative disease; mitochondria.
XX	Human caspase activator Smac.
OS	Homo sapiens.
XX	US6110691-A.
PN	US6110691-A.
XX	29-AUG-2000.
PD	29-AUG-2000.
XX	

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

submitted to the EMBL Data Library, October 1995

A:Reference number: Z19700

A:Accession: T23157

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1133 <N12>

A:Cross-references: EMBL:Z66514; PIDN:CA91344.1; GSPDB:GN00021; CESP:F59A2.6

A:Experimental source: clone K01A11

C:Genetics:

A:Gene: CESP:F59A2.6

A:Map position: 3

A:Introns: 13/2; 43/3; 107/3; 413/3; 492/3; 567/3; 635/3; 710/3; 738/3; 795/2; 1008/3; 1

Alignment Scores:

Pred. No.:	0.197	Length:	1133
Score:	116.50	Matches:	54
Percent Similarity:	42.73%	Conservative:	43
Best Local Similarity:	23.79%	Mismatches:	75
Query Match:	4.76%	Indels:	55
DB:	2	Gaps:	9

US-09-939-293-1 (1-1358) x T22976 (1-1133)

```

QY 197 GCACAGAAATCAGAGCCCTCATTCCTTAGTAGTGAAGCATGATGAGAGACAGTGTCT 256
    |||||
DB 436 AAGAGLysArgSerSerGluGluGluGluThrAlaAsnGluMetValArgSerLeuThr 455
QY 257 TTGGTACAGATAGACCTTAC----- 280
    |||||
DB 456 AAThrLeuGluAsnSerAsnSerGluThrGluLeuLeuLysGluLysLeuThrLeu 475
    |||||
QY 281 -----TTTCTCTCAGACACATATGCTTATGACCTTACT--- 322
    |||||
DB 476 AspLysGluLeuGluAlaArgGlnGlnThrGluLysAlaLeuThrGluLysGluLysVal 495
    |||||
QY 333 -----GAATATCTAAGCTGTTTATCTTACT---ACT 352
    |||||
DB 496 LeuThrThrSerLeuValGluLysGluGluGlnThrAlaGlnGlnGlnGlnGlnThr 515
    |||||
QY 353 TCTCTTACCGCAATATATCAAGTTTACTGGAAATGATATTCAGAGAGAGAGATGAA 412
    |||||
DB 516 GlnIleTyrGlnMet-----GluValGluLysGluLysValGlu 529
    |||||
QY 413 GGTGTGGAGGTATCATGAGACGAGAGCTGAGTACCTTCAACACAGAGTACTG 472
    |||||
DB 530 LeuValLysValGlnLeuGlnGlnAlaAlaGlnSerSerSerAlaGluGluAlaLeu 549
    |||||
QY 473 AAGCTGGAACCACTTGATGACTGCA-----GTTGCTTTTCAAGAGATGACAGAGAA 526
    |||||
DB 550 ArgAlaGluLysGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 569
    |||||
QY 527 GCTGCATATCAAACTGGCCGAGATCAGGCTTATACCCGACAGATAC----- 577
    |||||
DB 570 AlaLeuAsnSerLeuLeuLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 589
    |||||
QY 578 -----ATTCAGCTGTGTAAGCTCAGGTGGAAGAGGAGGAGCCAGCTC 619
    |||||
DB 590 ValGluLysGluGlnLysGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 609
    |||||
QY 620 TCCCGGAAAGCAAACTGGCAGAGACAGACAGATGAAGAGCTCCGTCAGAAAGA 679
    |||||
DB 610 SerSerSerValGluGlnGlnAlaLeu---ArgAlaGlnLysGlnLysGlnLysGln 628
    |||||
QY 680 CAGGAGGAAGGAGGAGGAGG-----GCTAGCTGCGAGAG 715
    |||||
DB 629 GlnGluLysGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 648
    |||||
QY 716 GAG-----GCTTACTGCTGAG 733
    |||||
DB 649 GlnThrAlaGlnLysGlnLys 655
    |||||

```

RESULT 4
S02041

dystrophin, muscle - chicken
N:Alternate names: duchenne muscular dystrophy protein
C:Species: Gallus gallus (chicken)
C:Date: 07-Sep-1990 #sequence, revision 27-Jun-1994 #text, change 16-Jul-1999
C:Accession: S02041; S02013; S71487
R:Lemaire, C.; Heilig, R.; Mandel, J.L.
Nucleic Acids Res. 16, 11815-11816, 1988
A:Title: Nucleotide sequence of chicken dystrophin cDNA.
A:Reference number: S02041; MUID:89098331; PMID:3062582
A:Accession: S02041
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-3660 <LEM>
A:Cross-references: EMBL:X13369; NID:963369; PIDN:CA91344.1; PID:963370
A:Note: 1869-His, 1885-Arg, and sequences lacking 1171-Met were also found
R:Lemaire, C.; Heilig, R.; Mandel, J.L.
EMBO J. 7, 4157-4162, 1988
A:Title: The chicken dystrophin cDNA: striking conservation of the C-terminal coding
A:Reference number: S02013; MUID:89210800; PMID:3072195
A:Accession: S02013
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-3573; 'HA', 3576-3660 <LEM2>
R:Heilig, R.; Lemaire, C.; Mandel, J.L.
Nucleic Acids Res. 15, 9129-9142, 1987
A:Title: A 230kb cosmid walk in the Duchenne muscular dystrophy gene: detection
A:Reference number: S09071; MUID:88067745; PMID:2825128
A:Accession: S71487
A:Molecule type: DNA
A:Residues: 222-281 <HEI>
C:Comment: Dystrophin is proposed to play a role in anchoring the cytoskeleton to ti
C:Comment: Defects in dystrophin are responsible for the Duchenne/Becker muscular dy
C:Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dy
C:Keywords: actin binding; calmodulin binding; cytoskeleton; leucine zipper; membra
F:18-237/Domain: alpha-actinin actin-binding domain homology <ACT>
F:253-327/Region: hinge
F:340-449/Domain: spectrin/dystrophin repeat homology <SP01>
F:450-558/Domain: spectrin/dystrophin repeat homology <SP02>
F:560-669/Domain: spectrin/dystrophin repeat homology <SP03>
F:670-719/Region: hinge
F:720-830/Domain: spectrin/dystrophin repeat homology <SP04>
F:838-936/Domain: spectrin/dystrophin repeat homology <SP05>
F:940-1047/Domain: spectrin/dystrophin repeat homology <SP06>
F:1049-1156/Domain: spectrin/dystrophin repeat homology <SP07>
F:1157-1265/Domain: spectrin/dystrophin repeat homology <SP08>
F:1267-1369/Domain: spectrin/dystrophin repeat homology <SP09>
F:1376-1479/Domain: spectrin/dystrophin repeat homology <SP10>
F:1480-1570/Domain: spectrin/dystrophin repeat homology <SP11>
F:1572-1678/Domain: spectrin/dystrophin repeat homology <SP12>
F:1680-1784/Domain: spectrin/dystrophin repeat homology <SP13>
F:1787-1877/Domain: spectrin/dystrophin repeat homology <SP14>
F:1878-1984/Domain: spectrin/dystrophin repeat homology <SP15>
F:1986-2103/Domain: spectrin/dystrophin repeat homology <SP16>
F:2105-2211/Domain: spectrin/dystrophin repeat homology <SP17>
F:2213-2319/Domain: spectrin/dystrophin repeat homology <SP18>
F:2323-2419/Domain: spectrin/dystrophin repeat homology <SP19>
F:2420-2467/Region: hinge
F:2468-2574/Domain: spectrin/dystrophin repeat homology <SP20>
F:2576-2683/Domain: spectrin/dystrophin repeat homology <SP21>
F:2685-2799/Domain: spectrin/dystrophin repeat homology <SP22>
F:2801-2928/Domain: spectrin/dystrophin repeat homology <SP23>
F:2930-3037/Domain: spectrin/dystrophin repeat homology <SP24>
F:3038-3075/Region: hinge
F:3052-3089/Domain: MW repeat homology <MW1>
F:3079-3357/Region: cysteine-rich
F:3481-3502/Region: leucine zipper motif
F:3547-3568/Region: leucine zipper motif

Alignment Scores:

Pred. No.:	0.524	Length:	3660
Score:	112.00 <td>Matches:</td> <td>56</td>	Matches:	56
Percent Similarity:	40.54% <td>Conservative:</td> <td>49</td>	Conservative:	49
Best Local Similarity:	21.62% <td>Mismatches:</td> <td>112</td>	Mismatches:	112

BEST AVAILABLE COPY
THIS PAGE BLANK (USPTO)

Query Match: 4.588 Indels: 42
DB: 1 Gaps: 8

US-09-939-293-1 (1-1358) x S02041 (1-3660)

QY 5 TCCGGCCCTGACAAAGGCGCTGTAAGAGTTGGCTGCCGACGCTACTTCATTC 64
DB 1163 SerleuA9gLVsAspLeuSerleuMetHsGluThrGlnAlaGluGluThr 1182
QY 65 TTCAGGTACAGACAGCTGTTGTGTCTTCCTGTTGCTTACCTTAAAGAGCGGTTC 124
DB 1183 LeuGlu-----ArgAspPheGluThrLeuThrPro 1192
QY 125 TCAGATTGATAGACATGCGACAAACTGTGACGATTGGCTTGGAGTAACCTGTGT 184
DB 1193 GluGluLeuGluLys----- 1197
QY 185 GCGGTTCTTATTCAGCAAAATCAGAGCCCTCATTCCTTAGTACGATGATGAGC 244
DB 1198 AlaValGluGluLeuLysArgAlaLysGluLysAlaMetGlnLysGluVal----- 1214
QY 245 AGACAGCTGTTTGTGTAACAGATAGACCTTACCTTCTCTCAGACCAATATGCG 304
DB 1215 --LysValLysLeuIleThrAspSerValAsnAsnPhelAlaLysAlaProProAla 1233
QY 305 TTGATTGACCTATTACTGATATATACAGGCTGTTTATACCTTACTTCTTACCGA 364
DB 1234 AlaAsnGluAlaLeuLysLysGluLeuAspValLeu-----IleThrSerTyrGlnArg 1251
QY 365 CAATATACACTTACTTGGGAAATGATTCAGAGAGAGAAATGAGTGTGCGAGGTG 424
DB 1252 LeuYsSerArgLeuAsnGlyLysCysLysThrLeuGlu-----GluValTyr----- 1267
QY 425 ATCATAGAGGACAGAGCTGATGATGATCTCAAAACACCAAGAGTACTTGAAGTGAACC 484
DB 1268 -----AlaCysTrpPheSerLeuLeuLeuSer-----TyrLeuAspAlaGluAsn 1281
QY 485 ACTTGATGACTGCTGCTGCTTTCAGAGATGCGACAGAGAACTGCTATCAACTGCG 544
DB 1282 LysTrpLeuAsnGluValGluLeuLysLeuLysAlaThrGlu-----AsnIleGlnGly 1300
QY 545 GCAATACAGGCTTATTAACCGCCAGAGATCATCATGAGTGTGAACTGAGGTGGA 604
DB 1301 AlaGluGluIleSerGluSerLeuAspSerLeuGluArgLeuMetArgHisProGluAsp 1320
QY 605 GAGGTGACACAGCTCTCCCGAAAGCAAAACCAAGCTGCGACAGACACAGATGAAGAG 664
DB 1321 AsnArgAsnGlnIleArgGluLeuAlaGlnThrLeuThrAspGlyLysLeuAspGlu 1340
QY 665 CTCCTGACAAACACAGAGAGAGAGAGAGCGGCTGCTGCGAGAGAGGCGCTAC 724
DB 1341 LeuIleAsnGluLysLeuGluLysPheAsnThrArgTrpGluGluGlnGlnGluAla 1360
QY 725 CTCGCTGAGGATTCAGGCGCTGAGACACAGCTGCTCTCCCACTGAGTGGGAAA 781
DB 1361 ValArgTrpGlnLysSerLeuGlnGlnSerIleGlnSerIleGlnGluThrAspLys 1379

RESULT 5

178557
N-methyl-D-aspartate receptor chain NMDAR2D-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence, revision 02-Aug-1996 #text, change 21-Jan-2000
C:Accession: 178557; 158158; D45219
R:Monyer, H.; Burnashev, N.; Laurie, D.J.; Sakmann, B.; Seeburg, P.H.
Neuron 12, 529-540, 1994
A:Title: Developmental and regional expression in the rat brain and functional properties
A:Reference number: 158158; MUID:94206533; PMID:7512349
A:Accession: 178557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1323 <RES>
A:Cross-References: GB:L31612; NID:9469068; PIDN:AAC37647.1; PID:9469069
A:Accession: 158158

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-66 'V', 68-1323 <RES>
A:Cross-References: GB:L31611; NID:9469066; PIDN:AAC37646.1; PID:9469067
R:Shi, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.; Kadotani, H.; Yokoi, M.; Akai
J. Biol. Chem. 268, 2836-2843, 1993
A:Title: Molecular characterization of the family of the N-methyl-D-aspartate receptor
A:Reference number: A45219; MUID:93155102; PMID:8428958
A:Accession: D45219
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1265-1323 <ISH>
A:Experimental source: Brain
A:Note: sequence extracted from NCBI backbone (NCBI:124265)
C:Superfamily: N-methyl-D-aspartate receptor 2D: glutamate receptor homology
F:451-879/Domain: glutamate receptor homology <GRH>

Alignment Scores:

Pred. No.:	0.569	Length:	1323
Score:	111.00	Matches:	56
Percent Similarity:	31.37%	Conservative:	8
Best Local Similarity:	27.45%	Mismatches:	73
Query Match:	4.60%	Indels:	67
DB:	2	Gaps:	9

US-09-939-293-1 (1-1358) x 178557 (1-1323)

QY 915 GCAAGCTGAGACGACAGGACACTCAGAGCTCAAAAGCGTCTGCGTATGGCA 856
DB 1004 AlaIleValArgGluGlnGluThrProThrLupProProAlaGlyAlaPheProGlyPhePro 1023
QY 855 GGGCAGACGTCGCGCTCTCTGCGTGCAGACAGATGATCCAAACCTGGCGAGGTG 796
DB 1024 SerProProAlaProProAlaAlaAlaAlaValGlyProProLeuLysArgLeu 1043
QY 795 GCA-----TCGCGCTGCTTCCCACTAGTGG----- 766
DB 1044 AlaPheGluAspGluSerProProAlaProSerArgTrpProArgSerAspProGluSer 1063
QY 765 -----GGAGACAGGACAGTGTCTCAGGCGCTCAATCTCCAGCAGGTAGGCC 718
DB 1064 GlnProLeuLysGlyGlyAla-----GlyGlyProSerAlaGly-AlaProThrAl 1081
QY 717 TCGTCTCCGACTGACCCGCTCTCC-----CTTCTCTGTGT 676
DB 1081 aProProArgArgAlaAlaProProProCysAlaTyrLeuAspLeuGluProSerPr 1101
QY 675 TTCAGAC-----GAGCTCTTATCTGTGCTTGTCCAGC 640
DB 1101 oSerAspSerGluAspSerGluSerLeuGlyGlyAlaSerLeuGlyGlyLeuGluProTr 1121
QY 639 TTGCTTCTG-----CTTCCGAGAGAGCTGTCACCT----- 606
DB 1121 pTrpPheAlaAspPheProTrpProTyrAlaGluArgLeuGlyProProProGlyArgTyr 1141
QY 605 -----CTTCCAGCTGC 595
DB 1141 rTrpSerValAspLysLeuGlyGlyTrpArgAlaGlySerTrpAspTrpLeuProProArg 1161
QY 594 AGTTTACCAAGCT-----GAATGTGATTCCTGCGGTTATAGAGGCTGTCT 547
DB 1161 gGlyGlyProAlaTrpHisCysArgHisCysAlaSerLeuGluLeuLeuProProArg 1181
QY 546 GCGCAGATTGATGACAGCTTGTGCTGCATCTGAAAGACCACTGAGTATGCCAA 487
DB 1181 gHis-----LeuSerCysSerHisAspLysLeuAspLysGlyGlyTr 1194
QY 486 GTGGTTTCCA 477
DB 1194 pTrpAlaPro 1197
RESULT 6
C45219

BEST AVAILABLE COPY
THIS PAGE